

## Neurons with Two Sites of Synaptic Integration Learn Invariant Representations

Konrad P. Körding

*koerding@ini.phys.ethz.ch*

Peter König

*PeterK@ini.phys.ethz.ch*

*Institute of Neuroinformatics, ETH/University Zürich, 8057 Zürich, Switzerland*

Neurons in mammalian cerebral cortex combine specific responses with respect to some stimulus features with invariant responses to other stimulus features. For example, in primary visual cortex, complex cells code for orientation of a contour but ignore its position to a certain degree. In higher areas, such as the inferotemporal cortex, translation-invariant, rotation-invariant, and even view point-invariant responses can be observed. Such properties are of obvious interest to artificial systems performing tasks like pattern recognition. It remains to be resolved how such response properties develop in biological systems. Here we present an unsupervised learning rule that addresses this problem. It is based on a neuron model with two sites of synaptic integration, allowing qualitatively different effects of input to basal and apical dendritic trees, respectively. Without supervision, the system learns to extract invariance properties using temporal or spatial continuity of stimuli. Furthermore, top-down information can be smoothly integrated in the same framework. Thus, this model lends a physiological implementation to approaches of unsupervised learning of invariant-response properties.

### 1 Introduction

---

**1.1 Invariant Response Properties in the Visual System.** The textbook view of the mammalian visual system describes a series of processing steps (Hubel & Wiesel, 1998). Starting with light-sensitive photoreceptors in the retina, signals are relayed by lateral geniculate nucleus and primary visual cortex toward higher visual areas. Along this pathway, neurons acquire increasingly complex response properties. Whereas an appropriately placed dot of light on a dark background is sufficient to strongly activate neurons in the lateral geniculate nucleus, cells in primary visual cortex respond best to elongated stimuli and oriented edges. At higher levels, complex arrangements of features are the optimal stimuli (Kobatake & Tanaka, 1994; Wang, Tanaka, & Tanifuji, 1996). Finally, in inferotemporal cortex of monkeys, neurons are tuned to rather complex objects, like faces or toys the monkey subject played with (Perrett et al., 1991; Rolls, 1992; Booth & Rolls, 1998).

However, a description of the visual system as a hierarchy of more and more complex filters is incomplete. In parallel to the increasing sophistication of receptive field properties, other aspects of the visual stimulus cease to affect firing rates (Rolls & Treves, 1997). In primary visual cortex, one major neuron type, simple cells, is highly sensitive to the contrast of the stimulus; if an oriented edge is effectively activating a neuron, the contrast-reversed stimulus usually is not (Hubel & Wiesel, 1962). Other neurons, complex cells, show a qualitatively different behavior. If an oriented edge is effectively activating a neuron, the contrast-reversed or phase-changed stimulus is usually equally efficient. Thus, the response of complex neurons is invariant with respect to the phase polarity of contrast of the stimulus (Hubel & Wiesel, 1962). Along similar lines, neurons in inferotemporal cortex show some invariance with respect to translation, scaling, rotations, and changes in contrast (Rolls, 1992). An even more extreme combination of specificity and invariance can be found in premotor cortex. Neurons may respond with high specificity to a stimulus, irrespective of its' being heard, seen, or felt (Graziano & Gross, 1998). Thus, a highly specific response to one variable, the position, is combined with invariance with respect to another variable, modality. The higher an area is located in the cortical hierarchy (Felleman & Van Essen, 1991), the more complex its neurons' receptive field properties are. Simultaneously, translation, scaling, and viewpoint invariance are more pronounced.

**1.2 The Computational Role of Invariances.** As invariant response properties are such a ubiquitous property of sensory systems, what are their computational advantages? In many categorization tasks, the output should be unchanged—or invariant—when the input is subject to various transformations. An important example is the classification of objects in two-dimensional images. A particular object should be assigned the same classification even if it is rotated, translated, or scaled within the image (Bishop, 1995). Invariant-response properties are especially important because they counteract the combinatorial explosion problem: highly specific neuronal responses imply small receptive fields in stimulus space. For a complete coverage of stimulus space, vast numbers of units are needed. In fact, the number of representative elements needed rises exponentially with the dimensionality of the problem. By obtaining invariant responses to some variables, receptive fields are enlarged in the respective dimensions, and the total number of neurons needed to describe a stimulus stays manageable.

Systems obtain invariant representations by several means (Barnard & Casasent, 1991). First, appropriate preprocessing can supply a neuronal network with invariant input data. Another option is to design the structure of the system so that invariances gradually increase. Finally, a system can learn invariances from the presented stimuli following principles of supervised or unsupervised learning.

**1.3 Invariances by Preprocessing.** Due to the advantages they offer to recognition systems, these invariances are frequently used in preprocessing for neural systems (Bishop, 1995). Applying a Fourier transformation and discarding the phase information produces data that are invariant with respect to translations of whole image. Even transformations that generate translation-, scaling-, and rotation-invariant representations are used (e.g., Bradski & Grossberg, 1995). Alternatively, the input might be scaled and translated before being processed by a neuronal network. As the generation of invariances by preprocessing is explicit, it is guaranteed that the network generalizes over the desired dimensions. They thus present a form of a priori knowledge about the kind of problems encountered. With this information, the network performs better, faster, and more reliably. However, this approach is limited to operations that can be explicitly defined. For example, it seems extremely difficult, if not impossible, to implement viewpoint invariance in this fashion.

**1.4 Mixing Invariance Generation with Processing.** In one class of networks, called weight-sharing systems, both processes—the generation of invariances and the feature extraction—are done step by step in the same network (Fukushima, 1980, 1988, 1999; Le Cun et al., 1989; Riesenhuber & Poggio, 1999). In the well-known Neocognitron (Fukushima, 1988), neurons at one level of the hierarchy have identical receptive fields but are translated in the visual field. Combining such weight sharing with converging connections leads to response properties at the next level that are somewhat translation invariant. The main drawback is that, similar to the case of generation of invariances by preprocessing, the constructor of the network needs to specify explicitly the desired invariances and thus to have specific a priori knowledge about which variables the processing is supposed to be invariant of.

**1.5 Supervised Learning of Invariances.** The obvious way to avoid the need to put in such specific a priori knowledge is the use of learning algorithms. Here, supervised and unsupervised approaches have to be differentiated. The former need labeled training data, and for large networks, a lot of these. With such data at hand, it is possible to learn invariant recognition, for example, training the network with a variant of the backpropagation algorithm (Hinton, 1987). To alleviate the problem of getting enough training data, the training set can be enlarged by applying appropriate transformations to individual examples. However, then we are back with an a priori specification of the invariance operation.

**1.6 Unsupervised Learning of Invariances.** Following this observation, several researchers started to investigate unsupervised learning of invariances. At this point we have to ask, When should two different stimuli be classified as instantiations of the same “real” thing? Let us consider a vi-

sual system for object recognition acting in the real world. A few principles about our world seem obvious:

1. If an object is present right now, it is likely to be present the next moment.
2. If an object covers a certain part of the retina, it is likely to cover its vicinity.
3. Objects tend to influence several modalities.

Several algorithms have been put forward for learning invariances in neural systems corresponding to these principles. Most notable are studies where variables are extracted from the input that smoothly vary in time (principle 1, proposed by Hinton, 1989; Földiak, 1991; Stone & Bray, 1995) or space (principle 2, Becker & Hinton, 1992; Stone & Bray, 1995; Phillips, Kay, & Smyth, 1995, cf. Becker, 1996). Principle 3 has been used by de Sa and Ballard (1998), but also is often considered a special case of principle 2, for example, auditory, visual, and somatosensory systems all allow a spatial localization. Still, this principle is more general and could enhance learning further. These systems can be compared to the networks described above in mixing invariances with processing. They share the advantages of needing few weights and combine them with the advantage of being able to learn those invariances in a way optimal to the task.

However, these mechanisms do not seem to map straightforward on biological principles (see section 4). In this article, we address this issue. First, we summarize recent electrophysiological results. Based on these, we define a suitably abstracted yet realistic learning rule. And finally, we demonstrate that it allows unsupervised learning of invariances exploiting spatial as well as temporal continuity. Thus, this learning rule is able to lend a physiological implementation to the algorithms described above.

**1.7 Relevant Physiological Results.** The most abundant type of neuron in cerebral cortex, the pyramidal cell, is characterized by its prominent apical dendrite. Recent research on the properties of layer V pyramidal neurons suggests that the apical dendrite acts, in addition to the soma, as a second site of synaptic integration (Larkum, Zhu, & Sakmann, 1999; Körding & König, 2000a). Each site integrates input from a set of synapses defined by their anatomical position and is able to generate regenerative potentials (Schiller, Schiller, Stuart, & Sakmann, 1997). The two sites exchange information in well-characterized ways (see Figure 1A). First, signals originating at the soma are transmitted to the apical dendrite by actively backpropagating dendritic action potentials (see Figures 1A and 1B; Amitai, Friedman, Connors, & Gutnick, 1993; Stuart & Sakmann, 1994; Buzsaki & Kandel, 1998) or passive current flow. Second, signals from the apical dendrite to the

soma are sent via actively propagating slow regenerative calcium spikes (see Figures 1D and 1E), which have been observed *in vitro* (Schiller et al., 1997) and *in vivo* (Hirsch, Alonso, & Reid, 1995; Helmchen, Svoboda, Denk, & Tank, 1999). These calcium spikes are initiated in the apical dendrites and cause a strong and prolonged depolarization, typically leading to bursts of action potentials (see Figures 1D and 1E; Stuart, Schiller, & Sakmann, 1997; Larkum, Zhu, & Sakmann, 1999). Experimental studies support the view that excitation to the apical dendrite is strongly attenuated on its way to the soma unless calcium spikes are induced (see Figure 1A; Schiller et al., 1997; Stuart & Spruston, 1998; Larkum, Zhu, & Sakmann, 1999). In conclusion, a subset of synapses on the apical dendrite is able to induce rare discrete events of strong, prolonged depolarization combined with bursts.

Experiments on hippocampal slices by Pike, Meredith, Olding, and Paulsen (1999) support the idea that postsynaptic bursting is essential for the induction of long-term potentiation. Furthermore, independent experiments support the idea that strong postsynaptic activity is necessary for induction of Hebbian learning (Artola, Bröcher, & Singer, 1990; Dudek & Bear, 1991). Integrating this with research on apical dendrites, we conclude that whenever the apical dendrite receives strong activation, calcium spikes are induced and synapses are modified according to a Hebbian rule (Hebb, 1949).

The generation of calcium spikes is very sensitive to local inhibitory activity. Even the activity of a single inhibitory neuron can effectively block calcium spikes (Larkum, Zhu, & Sakmann, 1999; Larkum, Kaiser, & Sakmann, 1999). Thus, it seems reasonable to assume that the number of neurons generating calcium spikes on presentation of a stimulus is limited on the scale of tangential inhibitory interactions. Here we allow only one neuron per stream and layer to feature a calcium spike in one iteration.

To complete the picture, we have to consider which afferents are targeting the apical and basal dendritic tree. The anatomy of a cortical column is complicated; nevertheless, some regular patterns can be discerned. The apical dendrites of the layer 5 pyramidal cells receive local inhibitory projections and long-range cortico-cortical projections (Zeki & Shipp, 1988; Cauller & Connors, 1994). Top-down projections from areas higher in the hierarchy of the sensory system usually terminate in layer 1, where many apical tufts can be observed (cf. Salin & Bullier, 1995). This supports the idea that top-down connections from higher to lower areas preferentially terminate on the apical dendrites.

The basal dendrites of the considered neurons receive direct subcortical afferents (e.g., the koniocellular pathway in visual cortex) in addition to projections from layer 4 spiny stellate cells. These are the main recipients of afferents from sensory thalamus and areas lower in the cortical hierarchy. Therefore, we use the approximation that the bottom-up input targets the basal dendritic tree (see Figure 1F).

## 2 Methods

---

The whole simulation is described by this pseudocode: Every neuron is

**For all iterations**

Calculate stimuli.

Calculate activities  $A$  going from lower areas to higher areas.

Calculates dendritic potentials  $D$  going from higher areas to lower areas.

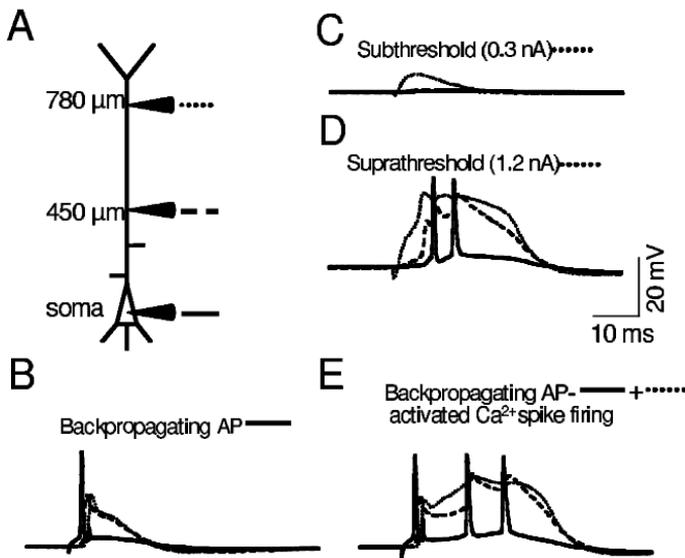
Determine in which neurons learning and thus calcium spikes are triggered.

Update the weights of those neurons according to Hebbian learning.

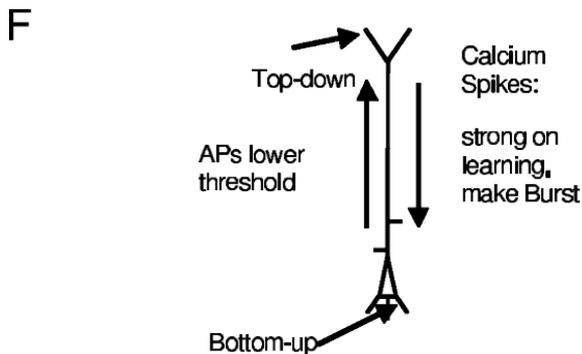
described by two main variables, corresponding to the two sites of integration (see Figure 1F):  $A$  is referred to as the activity of the neuron, and  $D$  represents the average potential at the apical dendrite. We simulate a rate coding neural network where a unit's output is a real number representing the average firing rate.

---

Figure 1: *Facing page*. The neuron model. (A) Experimental setup. A neuron is patched simultaneously at three positions: at the soma (recordings in solid line), 0.450 mm (recordings in dashed lines), and 0.780 mm (recordings in dotted lines) out on the apical dendrite. (B) Effect of inducing an action potential by current injection into the soma. The action potential does not only travel anterogradely into the axon but as well retrogradely into the apical dendrite, where it leads to a slow and delayed depolarization. (C) Injection of small currents at the apical dendrite does not trigger regenerative potentials and induces only a barely noticeable depolarization at the soma. (D) Injection of larger currents into the apical dendrite elicits a regenerative event with a long depolarization (dotted line). This slow potential is less attenuated and reaches the soma with a delay of only a few milliseconds (solid line). There, a series of action potentials is triggered, riding on the slow depolarization. (E) Combining a subthreshold current injection into the apical dendrite with a somatic action potential leads to a calcium spike and a burst of action potentials. Thus, somatic activity decreases the threshold for calcium spike induction. (F) Essential features incorporated into the model. Action potentials interact nonlinearly with calcium spike generation at the apical dendrite. Calcium spikes induce burst firing, gating plasticity of all active synapses. (In vitro data kindly supplied by M. E. Larkum (MPI für medizinische Forschung, Heidelberg), modified after a figure in Larkum, Zhu, & Sakmann, 1999.)



Modified after Larkum et. al. Nature 398:338-341



Depending on the layer, two different transfer functions are used that define the activity  $A_i^{(j)}$  of the neuron  $i$  in layer  $j$ :

$$\text{Sum: } A_i = \Theta \left( \text{sum} (\mathbf{A}_{\text{pre}} \cdot \mathbf{W}_i) - \langle A_j \rangle_j \right) / \left( N_{\text{pre}} \langle A_i \rangle_i^2 \right) \quad (2.1)$$

$$\text{Max: } A_i = \Theta \left( \text{max} (\mathbf{A}_{\text{pre}} \cdot \mathbf{W}_i) - \langle A_j \rangle_j \right) / \left( N_{\text{pre}} \langle A_i \rangle_i^2 \right), \quad (2.2)$$

where  $\Theta$  is the Heaviside function,  $\cdot$  is the element-wise product,  $\mathbf{A}_{\text{pre}}$  is the presynaptic activity vector,  $N_{\text{pre}}$  is the number of neurons presynaptic

to the basal dendrite,  $\langle A_j \rangle_j$  is the average activity within the layer, and  $\langle A_i \rangle_t$  is the running average of the neuron's activity with exponential decay and a time constant of 1000 iterations. Depending on the layer, the activity of a neuron results from the standard scalar product (sum case) or is determined by the most effective input only (max case). These types of activation functions have been discussed by Riesenhuber and Poggio (1999). The  $\langle A_j \rangle_j$  term mediates a linear inhibition. Summarizing, the neuron's activity is a linear threshold process with soft activity normalization and an activation function that is specific for each layer.

The apical potential  $D$  is calculated according to:

$$D_i = \mathbf{W}_i \mathbf{A}_{\text{pre}} + \alpha A_i \quad (2.3)$$

Due to the different connectivity to basal and apical dendrite, apical and basal dendrite have different  $\mathbf{A}_{\text{pre}}$  and thus also different  $\mathbf{W}_i$ ;  $\mathbf{A}_{\text{pre}}$  of the third-layer basal dendrite is 50-dimensional (number of second-layer neurons) and  $\mathbf{A}_{\text{pre}}$  of the apical dendrite is 4-dimensional (number of third-layer neurons).  $D$  has two components: the input from other neurons and an effect of the neuron itself.

For both the apical and the basal dendrites, the weight change is calculated as:

$$\begin{aligned} \Delta \mathbf{W}_i &= \eta^* (\mathbf{A}_{\text{pre}} + \mathbf{C}_{\text{pre}} - \mathbf{W}_i) \\ &\quad + \varphi^* (t/N_{i,\text{pre}} - 0.5) \quad : \text{if } D_i = \max(\mathbf{D}) \\ \Delta \mathbf{W}_i &= 0 \quad : \text{otherwise,} \end{aligned} \quad (2.4)$$

where  $\eta$  is the learning rate,  $N_{i,\text{pre}}$  the number of neurons presynaptic to the apical dendrite,  $t$  is the number of iterations since the neuron last learned, and  $\varphi$  is a constant. The  $\varphi$  term ensures that no neuron can stop learning and thus avoids the occurrence of so-called dead units.  $C = 1$  for the one presynaptic neuron with the highest  $D$ , and thus a calcium spike associated with learning, and 0 for the others. Only the neuron with highest  $D$  learns; neurons that do not learn for a large number of iterations increase their probability of learning.

Initially all weights are chosen randomly in the interval  $[0 \dots 1]$ . The default parameters are  $\eta = 0.002$ ,  $\varphi = 0.00005$ ,  $\alpha = 1$ . The system is simulated for 40,000 iterations unless stated differently.

All networks examined have three layers unless otherwise indicated. The network itself is organized into streams. Each stream receives input from a disjointed set of neurons. Only neurons of the highest layer receive contextual information to their apical dendrites by self-feedback or feedback from other areas. Otherwise the network is purely feedforward, with connections targeting the basal dendrites. Coupled layers are fully connected; there are no connections of any layer onto itself.

To explore different properties of the system, two sets of simulations were performed. In the first set (three-layer simulations), stimuli are rectangular similar to those used in physiological experiments. The second set (two-layer simulations) uses sparse, preprocessed input.

In the three-layer simulations, the first layer (input) consists of 10 by 10 neurons; the second layer of 50 neurons uses the SUM activation function; and the third layer of 4 neurons uses MAX. Stimuli resemble “bars” as used in physiological experiments. Their luminance has a gaussian profile orthogonal to the long axis with a length constant of 1. They are described by two parameters: the orientation of the bar and the position when projected on a line that is perpendicular to the bar’s main axis. The stimuli presented to each stream are correlated in orientation and position, as described in section 3.

We quantify the results of these simulations using different measures. For each stimulus, characterized by its orientation ( $\vartheta$ ) and position ( $r$ ), the average response is calculated during the second half of the simulation (iteration 20,000 to 40,000). Stimulus parameters  $\vartheta$  and  $r$ , which are drawn from a continuous distribution, are binned onto a  $20 \times 20$  grid covering the complete stimulus space. The responses to all stimuli that fall into the same bin are averaged and plotted in the  $\vartheta$ - $r$  diagram. Orientation- and position-specific receptive fields have a single peak in stimulus space, which drops along both dimensions. Neurons with translation-invariant responses are characterized by anisotropic  $\vartheta$ - $r$  diagrams.

A layer’s responses to orientation and position can be quantified by a bar specificity index calculated as follows. The  $\vartheta$ - $r$  diagrams are summed along the position axis, and the result is divided by its mean. Its standard deviation (over orientation) is averaged over all cells of the layer, yielding the orientation specificity  $\sigma(\text{orientation})$ . Position specificity  $\sigma(\text{position})$  is calculated analogously, summing along the orientation axis and calculating the standard deviation along the orientation axis.

To compare representations on the highest level in both streams, a cross-correlation coefficient CC of the activity patterns of both third layers is calculated:

$$CC = \frac{\sum_{ij} \langle (A_i^{(1)} A_j^{(2)})_t^2 \rangle}{\sqrt{(\sum_{ij} \langle (A_i^{(1)} A_j^{(1)})_t^2 \rangle) (\sum_{ij} \langle (A_i^{(2)} A_j^{(2)})_t^2 \rangle)}} \tag{2.5}$$

where superscripts denote the number of the stream and  $i, j$  in the sum run over all combinations, from 1 through 4, and  $\langle \rangle_t$  denotes the average over the considered iterations. This is a measure of the coherence of variables extracted by the two streams. If neurons in both streams have extracted identical variables, CC is 1.

For the control in Figure 3D, a special type of simulation is performed. The set of neurons with two sites of synaptic integration of the third layer is exchanged by a set of neurons with just one site of synaptic integration.

Both the second layer of the same stream and the third layer of the other stream are presynaptic to the same cell. We thus calculate the activities in the following iterative way:

$$I_{\text{input}}^{(i)}(0) = \max(\mathbf{A}_{\text{pre,same}} \cdot \mathbf{W}_i) \quad (2.6)$$

$$A_i(0) = \Theta(I_{\text{input}}^{(i)}(0) - \langle I_{\text{input}}^{(j)}(0) \rangle_j) / (N_{\text{pre}} \langle A_i \rangle_i^2) \quad (2.7)$$

$$I_{\text{input}}^{(i)}(n+1) = A_i(0) + m \cdot \mathbf{A}_{\text{pre,other}} \cdot \mathbf{W} \quad (2.8)$$

$$A_i(n+1) = \Theta(I_{\text{input}}^{(i)}(n+1) - \langle I_{\text{input}}^{(j)}(n+1) \rangle_j), \quad (2.9)$$

where  $n$  is the number of the iteration,  $\mathbf{A}_{\text{pre,same}}$  is the cell's input from the second layer,  $\mathbf{A}_{\text{pre,other}}$  the input from the other stream's third layer, and  $m$  is a scaling factor. Since  $A$  changes depending on the other stream's  $\mathbf{A}$  and vice versa, the activities change over subsequent updates. We use  $A_i(20)$  as an approximation of the convergence value. Twenty iterations are sufficient for the process to converge to values very near the final value (data not shown) unless the process diverges to infinite activities (at  $m$  of about 1). After this process, the neuron with the highest activity learns with the same  $\Delta \mathbf{W}$  described above.

In the two-layer simulations, the first layer consists of 12 neurons and the second layer of 4 neurons. The main difference from the three-layer simulations is that the inputs are already preprocessed; the input layer for the two-layer simulations can be compared to the second layer of the three-layer simulation. The stimuli are 2D maps of size  $4 \times 3$ . One axis is labeled class (4 neurons) and the other instantiation (3 neurons). To generate a stimulus, we first select a number of active classes, which are identical for all streams. The probability for  $n$  classes active at the same iteration is proportional to a parameter  $p_c$  to the power of  $n$ . Then for each class, we set an instantiation, which is chosen independently in all streams and thus can also be different. All of the chosen elements of the input map are set to 1, and the rest are set to 0. These stimuli have the following property: at low  $p_c$ , the stimuli are very similar to the ones of the three-layer simulations, class is perfectly correlated, and instantiation is not correlated. The higher  $p_c$ , the larger the spurious correlation in the class. It thus is a measure that makes the task difficult. We used this more abstract simulation to be able to do extensive explorations in larger systems with several streams.

To quantify properties of a whole layer of neurons, a class specificity index is calculated. If a class is assigned to a neuron, we can quantify how specific it is to that class by calculating its average activity for that class minus the average activity for all other classes:  $S = \langle A \rangle_{\text{class}} - \langle A \rangle_{\text{nonclass}}$ , where  $\langle \cdot \rangle_x$  is the average over a number of iterations under the condition

that the stimulus is chosen from  $x$ . The class specificity is defined as the maximum of that specificity measure over all neuron class assignments, where every class is assigned to exactly one neuron. If the class specificity is large, neurons are specific to one class and ignore the rest. It is small if several neurons code for the same class.

### 3 Results

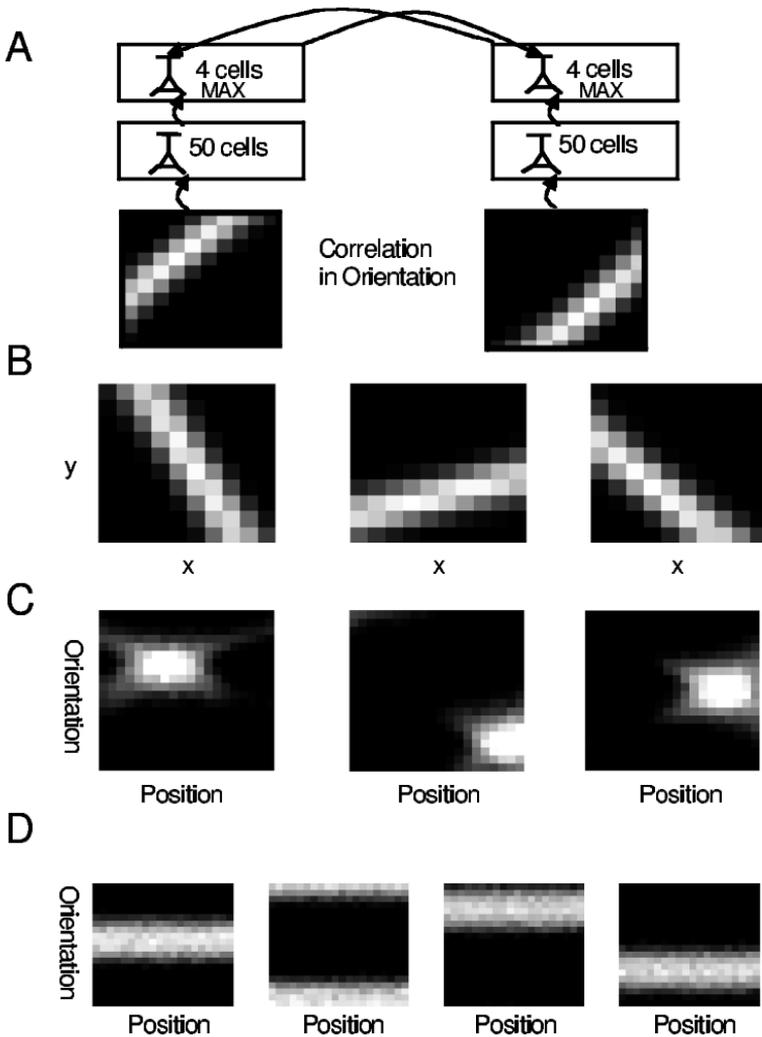
---

Several simulations are performed to characterize the system (see equations 2.1–2.4) and effects of different inputs on the basal and apical dendrite. We start with a system that implements an approximation to the spatial smoothness criterion. Next, we move on to a system that approximates the temporal smoothness criterion and explore the interaction with top-down signals. Finally, we show how lateral connectivity improves learning in the system.

**3.1 Spatial Continuity.** We explore a system with two input streams (see Figure 2A), which are treated as analyzing spatially neighboring but nonoverlapping parts of the visual scene. In this simulation, we first address principle 1 from above: the spatial continuity of stimuli. The input to each stream consists of rectangular stimuli characterized by their orientation and position. The orientations of the stimuli presented to each stream are perfectly correlated, whereas stimulus position is not correlated. As a control, these assumptions are relaxed in additional simulations, and the results are reported further down. The activity of neurons in the first layer is directly set by the luminance of the stimulus. At the third level, the output of each stream projects to the apical dendrites in the other stream.

Receptive fields of three neurons randomly chosen from the second layer are shown in Figure 2B and closely resemble individual input stimuli. They are position and orientation specific and thus localized in stimulus space (see Figure 2C). This can be understood from the absence of specific excitatory projections to the apical dendritic tree. As a consequence, calcium spikes are triggered by the depolarization induced by backpropagating action potentials only, and the learning rule applied is effectively equivalent to competitive Hebbian learning. The calcium spike dynamics here has the effect of ensuring a mechanism for competition. Inhibitory interactions in the network enforce that different neurons learn different stimuli (Körding & König, 2000b). Thus, these neurons tend to cover the input space uniformly.

Third-layer neurons show qualitatively different response properties (see Figure 2D). Each neuron responds selectively to stimuli of a single orientation, regardless of the position. Reciprocal projections between the two streams targeting the apical dendrites favor correlated stimuli to trigger calcium spikes. As stimulus orientation is correlated across streams, neurons extract this variable and discard information on the precise position. Thus,



the “supervision” of each stream by the other leads to invariant-response properties and approximates criterion 2.

**3.2 Controls.** To analyze the properties of the proposed system further, we investigated coverage of stimulus space by neurons in layers 2 and 3. Figure 3A demonstrates that the standard deviations of total activity in layers 2 and 3 are small compared to the mean (5.3% and 6.5% in layers 2 and 3, respectively). Thus, competition implemented by the inhibitory action on calcium bursts is sufficient in the second as well as in the third layer to guarantee an even coverage of stimulus space.

As a next step, we compare the effects of different learning rates (see Figure 3B). To quantify convergence, the coherence CC (see equation 2.5) of responses between streams is determined as described in section 2. A low learning rate ( $\eta = 0.0005$ ) leads to a slow convergence ( $\sim 13,500$  stimulus presentations until reaching  $CC = 0.75$ ), but CC reaches high levels, saturating at 0.96 (mean for the last fourth of the simulation) very close to the theoretical maximum of 1. When 4 and 16 times higher learning rates (the 4 times larger learning rate is used in the other simulations) are used, convergence is faster by a factor of 2 and 4 ( $\sim 7000$  and  $\sim 4000$  stimulus presentations, respectively). However, CC is lower than before (0.94 and 0.88 for 4 times and 16 times learning rates respectively). We want to state explicitly that the system converges rather slowly and theoretically it should be possible to learn invariances more quickly. Nevertheless, the system is robust enough to support extraction of coherent variables and development of invariant responses in a reasonable time.

An important question is how sensitive the model is with respect to mixing the bottom-up input into the lateral learning signal. To investigate this question, we systematically changed the ratio ( $\alpha$ ) of the learning signal determined by the stimulus to the part defined by the contextual input. In Figure 3C, it can be seen that at an  $\alpha$  of about 1, implying that the effect of the cell's own activity on learning is of the order of the effect of the context cells, invariance generation breaks down. This furthermore leads to an experimentally testable hypothesis. The lateral and the somatic effect on calcium spike generation can experimentally be distinguished.

It is necessary to control that the same results could not be obtained using a neuron model with just one site of synaptic integration. To test this, we constructed a simulation where the third-layer cells were exchanged by cells with just one site of synaptic integration (see equations 2.6–2.9). The relative contribution of the input from the other stream can be gauged by a scaling parameter  $m$ . At a scaling parameter  $m$  of about 1, the activities

---

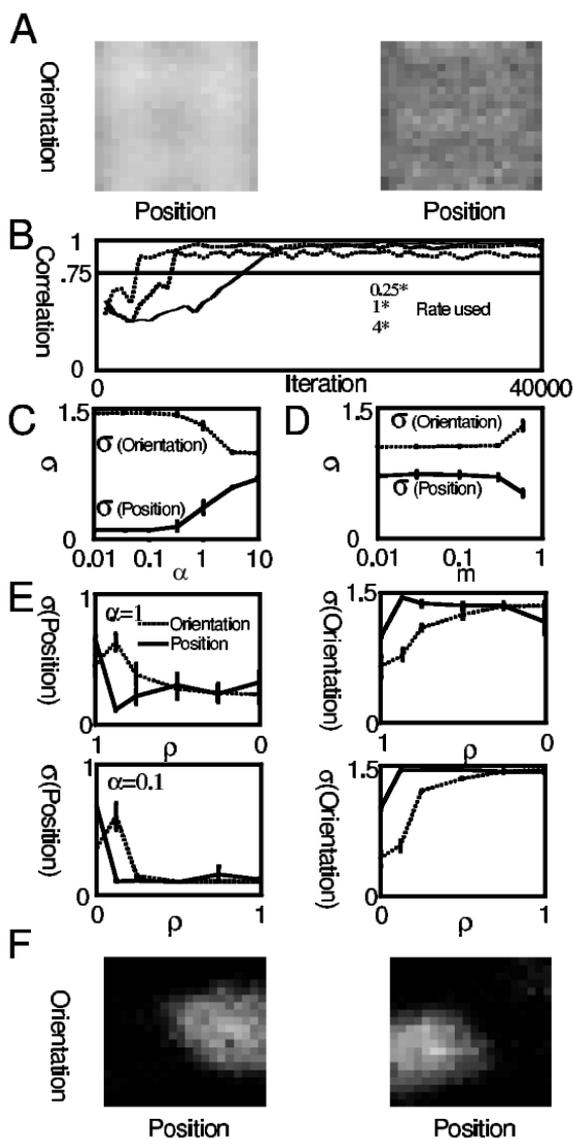
Figure 2: *Facing page*. Basic mode of processing. (A) The network consists of two separate streams. Only neurons on the highest layer have cross-stream connections. An example stimulus as given to the input layer is shown in gray scale. (B) The resulting receptive fields (columns of the weight matrix from the input layer) of three neurons in the second layer are shown gray scale coded. Light shades indicate sensitive regions; dark areas indicate insensitive spots. (C) Response strength (firing rate  $A$ ) as a function of stimulus position and orientation of layer 2 neurons. For stimuli with orientation given by the ordinate and position given by the abscissa, the activity of the corresponding neuron is shown on a gray scale (see section 2). The same gray scale is used for all three units. The neurons analyzed in C are identical to those in B. (D) The corresponding diagram as in C is shown for neurons in layer 3. Note the homogeneity of receptive fields along the horizontal axis, indicating translation-invariant response properties.

start to diverge, and no valid values can be obtained. Figure 3D shows that in the one-cell model for small  $m$ , the cells do not properly get invariant. For higher  $m$  of 0.6, the cells get more invariant, albeit still far less invariant than in the two-site model. But at  $m = 0.6$ , the cells are no longer really local detectors. Sixty percent of their activity is determined by the other stream. Then, however, neurons are no longer local feature detectors. In learning networks of the architecture investigated here, where neurons perform only one synaptic integration, a trade-off seems to exist: either neurons act as local detectors and do not properly learn invariant responses, or neurons are no longer local detectors but learn correctly. This is the reason that we argue that two sites of synaptic integration indeed improve processing.

In the simulations described above, we assumed a perfect correlation of the orientation of stimuli presented to the two streams and a complete lack of correlation of position. Indeed, in natural occurring stimuli, orientation is correlated much stronger than position (Betsch, Einhäuser, Körding, & König, 2001). Nevertheless, the assumptions made in the above simulation might be too strong. Therefore, we investigated performance of the learning rule while varying correlation of orientation and position. The orientation of the stimulus presented to the first stream ( $\vartheta_1$ ) is randomly selected in the interval  $[0 \dots \pi]$  as before. The orientation of the other stimulus is chosen equally distributed in the interval  $[\vartheta_1 - \pi(1 - \rho)/2 \dots \vartheta_1 + \pi(1 - \rho)/2]$ . The parameter  $\rho$  allows the correlation of stimulus orientation

---

Figure 3: *Facing page*. Controls: (A) The diagram shows total activity in layer 2 (left) and layer 3 (right) as a function of stimulus orientation and position. The small variations in shading indicate that the network activity is comparable for all stimuli. (B) The cross-stream correlation is shown (see section 2) for different learning rates. (C) The coupling of the two integration sites ( $\alpha$ ) is varied; measures for the bar specificity (see section 2) for position  $\sigma(\text{position})$  (solid line) and for orientation  $\sigma(\text{orientation})$  (dotted line) are plotted. (D) As a control, a simulation where the neurons of the third layer exhibit just one site of synaptic integration was performed (see section 2). The influence of lateral inputs to the cell is gauged by a parameter  $m$  (low  $m$  means low lateral effects). The orientation  $\sigma(\text{orientation})$  and position  $\sigma(\text{position})$  specificity is plotted against the influence parameter  $m$ . (E) As in C, the bar specificities  $\sigma(\text{position})$  and  $\sigma(\text{orientation})$  are plotted, changing the position of stimuli between both streams (left column) or changing orientation of stimuli between both streams (right column). The correlation of stimulus properties across the two streams ( $\rho$ ) is varied for two values of  $\alpha$  ( $\alpha = 1$ , strong coupling of integration sites and  $\alpha = 0.1$  medium coupling of integration sites). The solid and dotted lines show data of simulation runs where the correlation of stimulus position or orientation was varied respectively. In either case  $\rho = 0$  corresponds to the parameters used in the other simulations. (F) Response properties of layer 3 neurons are shown when identical stimuli presented to both streams.



to change smoothly, with  $\rho = 1$  resulting in perfect correlation as used in most simulations presented here and  $\rho = 0$  resulting in zero correlation. In Figure 3D, the dotted lines show that depending on the strength of coupling between the two sites  $\alpha$ , that is, when learning is determined to a large extent from the context, even big stimulus intervals ( $\rho = 0.25$ ) corresponding to a small correlation still allow learning of invariances.

Along similar lines, we analyzed the influence of varying degrees of correlations in the position on learning of invariances. Again we observe a robust behavior of the learning rule, tolerating significant variations of the parameter  $\rho$ , here describing the correlation of position. In fact, it turns out that at least for a strong coupling of the two integration sites ( $\alpha = 1$ ), a finite correlation of stimulus position is helpful in learning invariances. This can be understood by investigating dynamics of learning. Weak correlation of position leads to a symmetry breaking, and the system leaves metastable states more quickly (data not shown). Obviously if both variables are perfectly correlated, learning of invariances is no longer possible. Then neurons develop receptive fields specific in both space and orientation (see Figure 3E). Receptive fields are large, as they have to be, in order for the set of neurons to cover stimulus space. But more important, response properties of all neurons are selective for orientation as well as for position. Variables, which are correlated across streams, are considered relevant according to criterion 2. Uncorrelated variables induce invariant responses and, thus, any information pertaining to these feature dimensions is discarded.

**3.3 Temporal Continuity.** In the simulation above, we used a pseudo-random sequence of stimuli with the time constants of the relevant variables set to zero. Thus, no interaction between subsequent stimuli occurred. On the other hand, temporal continuity of objects in the real world allows generating invariant responses. Here we analyze a system operating on this principle consisting of one stream only (see Figure 4A). The time constant of the potential at the apical dendrite ( $D(t) = \sum(A(t-n)^*(1-1/\tau_D))^n$  where  $n$  starts from 0) is set to a finite value ( $\tau_D = 10$  iterations) and stimulus orientation changes with the same time constant ( $\mathcal{G}_{\text{new}} = (\mathcal{G}_{\text{old}} + 0.1*\pi*(\text{rand} - 0.5))\text{mod}\pi$ , where  $\text{rand}$  is uniformly drawn from  $[0..1[$ ). Figure 4B shows that neurons in the third layer obtain orientation-selective but position-invariant receptive fields. Thus, the effect of a finite-time constant is comparable to the “supervising” input by a second stream, and the simulation using two streams could equally be interpreted as the slow temporal dynamics unfolded in space.

**3.4 Top-Down Contextual Effects.** The simulations described exploit temporal or spatial regularities of stimuli. Thus, they represent a bottom-up approach of generation of invariances. As a next step, we explore whether top-down signals can exploit the same mechanism. Similar to the simulation above, a single processing stream is used (see Figure 5A). A set of 10 units is added, representing an independent source of information. Their response properties are specified to be position selective and orientation invariant, and the union of their receptive fields evenly covers stimulus space. Training the network with a pseudo-random sequence of stimuli as before leads to an interesting result. Receptive fields of neurons in layer 3 are specific with

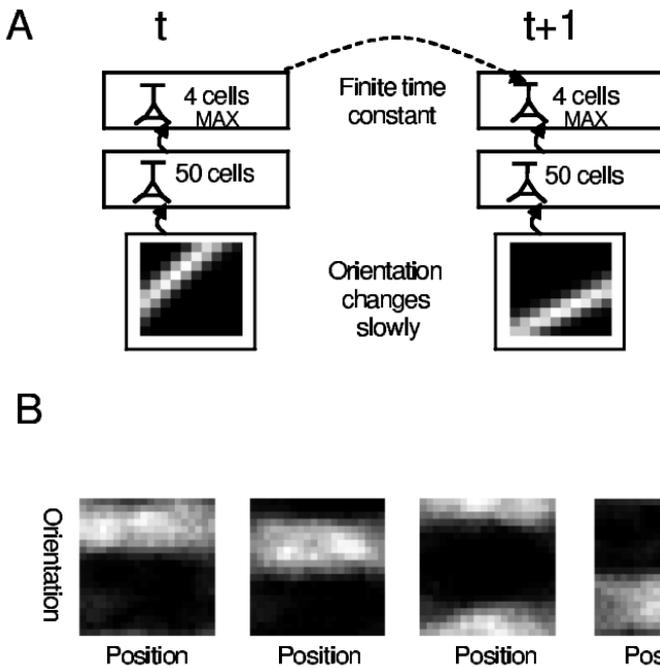
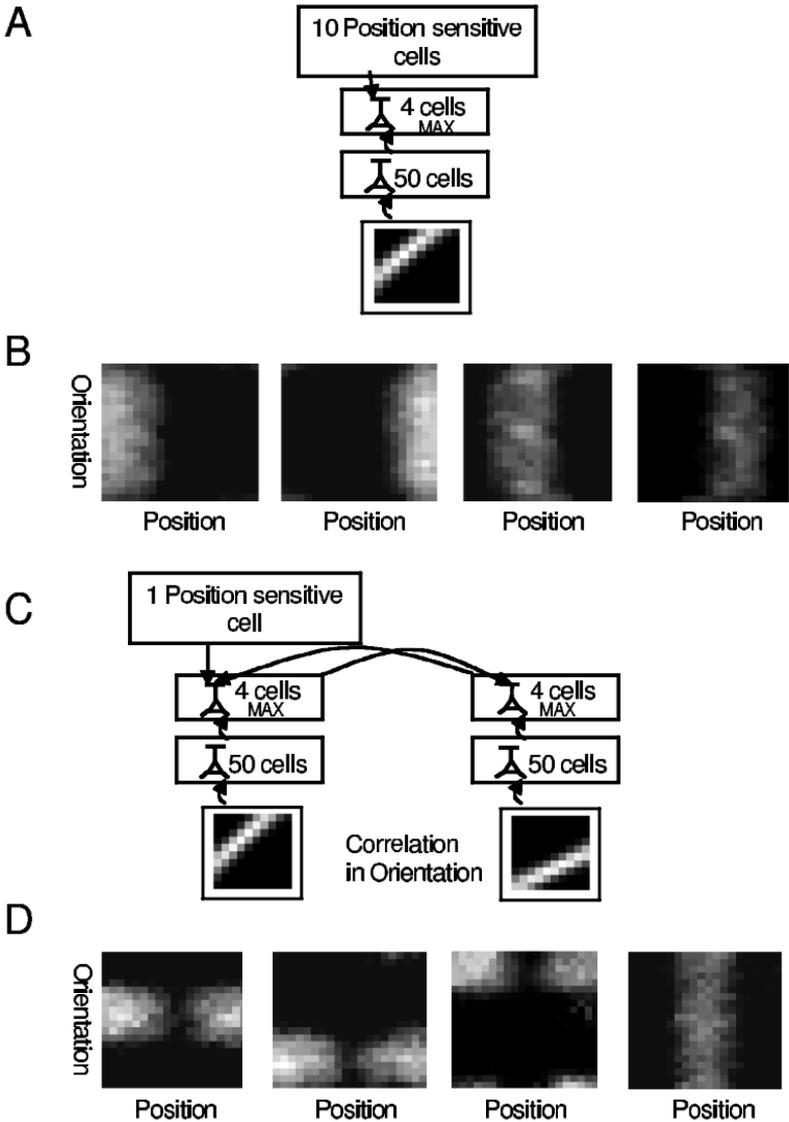


Figure 4: Learning from temporally smoothly varying patterns. (A) The network consists of one stream only. The apical dendrites have a finite time constant, and the orientation of the stimulus changes slowly. Note the similarities of the delayed interaction with the two-stream setup in Figure 3. (B) Response strength of layer 3 neurons shown as a function of stimulus orientation and position.

respect to position but invariant with respect to orientation (see Figure 5B). This demonstrates that the type of invariance is not solely defined by the stimulus set; the neurons learn to transmit the part of the information that is correlated with the activity of the other set of neurons. Thus, this represents a special case of the “maximization of relevant information” principle proposed in Körding and König (2000a).

As a next step a combination of top-down and bottom-up invariance extraction is investigated. The network consists of two coupled streams, one of these receiving an additional projection from one position-selective neuron (see Figure 5C). This neuron can be activated by stimuli of any orientation, but from just one-quarter of the positions. The network is trained with stimuli correlated in orientation but not in position. One neuron in the left module of the third layer learned to extract the information relevant for the position-selective neuron. The other neurons picked up the information correlated across streams (see Figure 5D). However, these two types of input



to the apical dendrite show some undesired interaction. Due to the strong competition implemented by the local inhibitory interaction on the calcium spikes, each stimulus is learned by one neuron only. As the top-down input induces one orientation-invariant, position-selective receptive field, this position is cut out of the receptive fields of the remaining orientation-selective, position-invariant neurons. Thus, they are actually only partly position invariant. It remains a major issue for future research how to extract several

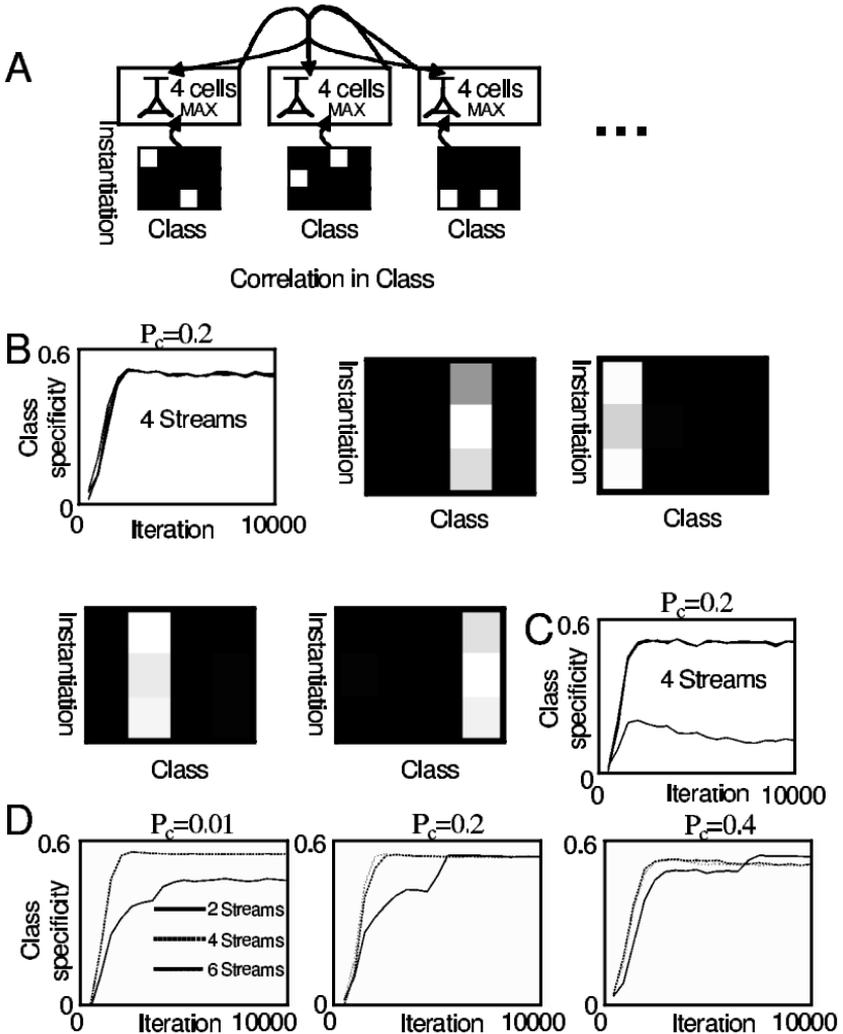
variables from several such principles simultaneously (Phillips et al., 1995). Nevertheless, given that we deliberately have chosen conflicting demands of top-down and bottom-up information, the network behaves reasonably well and shows that learning of invariance by spatial coherence can be complemented by top-down information on the relevance of different features. Apart from this obvious problem, this simulation demonstrates that lateral connections suffice to learn invariances and that these processes can be refined using top-down information.

**3.5 Multiple Streams.** In the simulations above, at most two streams were involved. This is, of course, a gross simplification; in the biological system, each cortical patch, analyzing part of the visual field, has many more neighbors. Therefore, we extend the network to include multiple streams and vary their number and the degree of correlation of inputs to these streams.

To allow more extensive simulations, we concentrate on the generation of invariant-receptive fields in the step from layer 2 to layer 3 in the previous simulations. Thus, we do not deal with the feature extraction from layer 1 to layer 2 but set activity levels of layer 2 neurons directly. The input is a preprocessed two-dimensional map. A stimulus consisting of a bar would activate mainly one layer 2 neuron. It is therefore represented by a binary pattern, with all entries set to zero but the one neuron with the corresponding orientation and position set to one. Presentation of several overlapping bars is coded accordingly. Because feature extraction is not dealt with in this simulation, we drop the terms *orientation* and *position* and use neutral descriptors of class and instantiation, respectively. In the simulation, instantiations (positions) are not correlated between different streams, and the correlation between classes (orientation) is regulated by the parameter  $p_c$ .  $p_c$  represents the amount of spurious correlations of orientation due to the presence of multiple oriented stimuli (see section 2). At small  $p_c$  in all streams, one unit of a class is activated. At larger  $p_c$ , spurious correlations between different classes are introduced, making the learning task harder. In natural scenes, typically a large number of objects and features is seen at

---

Figure 5: *Facing page*. Effect of the relevant infomax. (A) The network matches a single stream as shown in Figure 2 with another set of neurons added above the higher layer of that stream. (B) The resulting responses of third-layer neurons as a function of orientation and position are shown. (C) The second network is the same as in Figure 2 except for another set of neurons added above the higher layer of the left stream. The neurons on the third layer receive input from the third layer of the other stream and the highest layer. (D) Results with relevant infomax. The responses of the left stream's third-layer neurons are shown. The neurons on the right side develop position-selective and orientation-invariant response properties, whereas the other neurons respond orientation specific and translation invariant.



the same time. We therefore consider these simulations an important step on the way to being able to deal with natural images.

We use the class specificity measure (see section 2) to characterize the performance of the system. They converge to a rather high value and stabilize there (see Figure 6B for  $p_c = 0.2$ , 4 streams). Thus, typically neurons in all streams learn to represent one class. Occasionally it happens that two neurons in one stream do not code for a single class, but instead code for two classes each (see Figure 6C,  $p_c = 0.2$ , four streams). Such a metastable state can last for prolonged periods before typically collapsing into the correct solution (see Figure 6D, third panel). The convergence behavior and

receptive fields are shown for the typical case of Figure 6C. To analyze the behavior further, we varied  $p_c$  as 0.01, 0.2, and 0.4 and determined the effect of the number of streams (see Figure 6D). Increasing the number of streams speeds up convergence. The number of effective streams in visual cortex can be estimated by the ratio of extent of long-range connections to the minimal distance of neurons with nonoverlapping receptive fields. For the cat we obtain a (linear) ratio of about 7 mm / 2 mm = 3.5 (Salin & Bullier, 1995), giving an estimate ( $\pi r^2$ ) of 38 for the total number of streams. This high number places the cortical architecture in the parameter range where lateral connectivity allows much faster and more stable learning.

## 4 Discussion

---

We have shown that a reasonable neuron model with two sites of synaptic integration can be used to implement learning principles of spatial and temporal continuity. Furthermore, within the same framework, it can be combined with principles that approximate maximization of relevant information. Finally, lateral connections, as commonly seen in mammalian neocortex, speed up the learning process.

**4.1 Approximations.** In our implementation obviously several physiological aspects are simplified.

First, experiments on the physiological properties of dendrites of pyramidal neurons reveal a high degree of complexity (Johnston, Magee, Colbert, & Cristie, 1996; Segev & Rall, 1998; Helmchen et al., 1999). Some theoretical studies argue for a superlinear interaction of postsynaptic signals in the dendrite (Softky, 1994). Other studies imply linear summation (Bernander, Koch, & Douglas, 1994; Cash & Yuste, 1999), or provide evidence for sublinear dendritic properties (Mel, 1993). Furthermore, the effectivity of input via

---

Figure 6: *Facing page.* Multiple streams. (A) Each stream consists of two layers: an input layer where preprocessed sparse activity is present, projecting to the second layer. The modules of the second layer are connected to all other modules' second layer. The number of the streams is varied and is either 2, 4, and 6. (B) The left panel shows the class specificity index (see section 2) in the typical case as a function of the iteration. The other panels show the corresponding receptive fields for each neuron. (C) The class specificity is shown for a case where not every neuron was associated to exactly one class. (D) The class specificity is shown for each stream averaged over four runs. Each line represents one stream. The left-hand panel shows the average class specificity as a function of the iteration and the number of streams at  $p_c = 0.01$ . Note that the four-stream and the six-stream lines largely overlap. The middle panel shows that introducing spurious correlations ( $p_c = 0.2$ ) slows the convergence. This effect is more pronounced for a small number of streams. The right-hand panel ( $p_c = 0.4$ ) shows that at very strong false correlation, the effect also holds.

the apical dendritic tree is discussed. Obviously it is not possible to subscribe to all these views simultaneously. In a gross simplification, we assume linear summation of all inputs to the basal dendritic tree, and the function of the apical dendrite is reduced to a threshold mechanism triggering calcium spikes. This choice avoids extreme views on the unresolved matters described above and also avoids obscuring the article with too many details and parameters.

Second, in this article, we did not include an effect of calcium spikes on postsynaptic activity. However, it is known that calcium spikes can induce postsynaptic bursting activity (Larkum, Zhu, & Sakmann, 1999; Williams & Stuart, 1999; see Figure 1C). Within the framework of lateral and top-down projections terminating on the apical dendrite, this would result in contextual information not only gating learning, but also influencing signal processing itself (Phillips et al., 1995; Kay, Floreano, & Phillips, 1998). Indeed, it has been shown that top-down projections can enhance processing of sensory signals (Siegel, Körding, & König, 2000).

Third, the inhibitory neurons are reduced to a linear effect on somatic activity and act as mediators of the winner-take-all mechanism on calcium spikes. Larkum, Zhu, & Sakmann (1999) found a dramatic effect of inhibition on the calcium spike generation mechanism. The activity of a single inhibitory neuron seems to be sufficient to prevent triggering of calcium spikes. This effectively describes a winner-take-all mechanism that we implemented algorithmically. (For an alternative implementation, see Körding & König, 2000b.)

Fourth, in our simulations, we use a normalization scheme for the network activity. Actually, such a procedure is often used and may be implemented by superlinear inhibition. As currents mediated by GABA<sub>B</sub> receptors are observed at high presynaptic firing rates (Kim, Sanchez-Vives, & McCormick, 1997), this is a promising candidate mechanism.

Finally, in this work, a rate-coding system was implemented, ignoring the precise timing of action potentials. Indeed, in several of the experiments cited above (Markram, Lübke, Frotscher, & Sakmann, 1997; Larkum, Kaiser, & Sakmann, 1999) synaptic plasticity has been found to depend on the relative timing of afferent and backpropagating action potentials. Evidence is available that the mean firing rate of cortical neurons maps directly on the relative timing of action potentials (König et al., 1995). Thus, our approximation appears reasonable, and from existing knowledge, we can be optimistic that the effects observed here hold up in a more detailed simulation study.

**4.2 Comparison with Other Unsupervised Systems That Learn Invariances.** As discussed in section 1, several possible criteria can be used to learn invariances. The spatial and the temporal criteria are commonly used (Hinton, 1989; Földiak, 1991; Becker & Hinton, 1992; Stone & Bray, 1995; cf. Becker, 1996, 1999; Kay et al., 1998). Those studies at first sight do not seem to map directly on known physiology. We do not claim to outper-

form those studies with respect to convergence speed or the maximization of a goal function, but investigate physiological mechanisms that may be used by the brain to implement algorithms like these. Nevertheless, we do hope that investigating mechanisms used by the brain will finally help us conceive algorithms that outperform today's algorithms.

The main virtue of the learning rule proposed here is that it explicitly addresses the existence of a trade-off between correctly representing local stimulus features and correctly learning invariant representations. Furthermore, it allows smooth integration of top-down information with contextually guided feature extraction to result in extraction of relevant information. To learn invariant representations from natural stimuli is an important line of research, which is currently becoming feasible (Hyvarinen & Hoyer, 2000). O'Reilly and Johnson (1994) indicate that at least the temporal smoothness criterion could be implemented in one site of the synaptic integration model where delayed recurrent excitatory activity takes the role of supplying the signal necessary for learning. In their approach, the same trade-off is likely to hold; with weak delayed excitation, the learning rule effectively is Hebbian, and with strong excitation, neurons are not local feature detectors but have temporal low-pass characteristics. Eisele (1997) presents an interesting approach enhancing the temporal smoothness criterion to cases in which transitions are not symmetric. Two types of dendrites represent states a system can go to or can come from. These different kinds of properties are then mapped on apical and basal dendrites. Rao and Sejnowski (2000) propose a system where temporally predictive coding in a recurrent neural network leads to the generation of invariant-response properties. This model from our point of view uses a variant of the temporal smoothness criterion, and we assume it to be subject to the trade-off described above as well. Mel, Ruderma, & Archie (1998) investigate the development of complex receptive fields. Modeling cortical pyramidal neurons, bars of identical orientation are presented at different positions and Hebbian learning applied. Their approach thus needs a mechanism ensuring that neurons learn only when the same orientation is shown. Thus, a mechanism to gate learning is needed; our approach may be considered a physiological implementation. We must note, though, that due to its derivation from physiological results, it is not directly amenable to mathematical analysis, and thus it is not possible to write down its goal function explicitly. By virtue of the physiological implementation, we nevertheless obtain a unified framework allowing us to use temporal continuity and spatial continuity for the development of invariant-response properties and combine this with a bias to relevant information by top-down signals.

**4.3 Cortical Building Block.** Analyzing the mammalian neocortex shows an astonishing degree of similarity across areas as well across animal species. It is a shared feeling of many neuroscientists that with this anatomical similarity should go a common system for computing and learning. "The

typical wiring of the cortex, which is invariant irrespective of local functional specialization, must be the substrate of a special kind of operation which is typical for the cortical level" (Braitenberg, 1978). The principles proposed for supporting unsupervised learning of invariances should be part of such a common cortical building block; these principles in turn may shed some light on the significance of a prominent property of the architecture of cortical columns, the apical dendrite of pyramidal neurons.

## Acknowledgments

---

We thank Matthew E. Larkum for kindly supplying Figure 1. We also thank the Boehringer Ingelheim Fonds and the Swiss National Fonds for financial support.

## References

---

- Amitai, Y., Friedman, A., Connors, B. W., & Gutnick, M. J. (1993). Regenerative activity in apical dendrites of pyramidal cells in neocortex. *Cereb. Cortex*, *3*, 26–38.
- Artola, A., Bröcher, S., & Singer, W. (1990). Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature*, *347*, 69–72.
- Barnard, E., & Casasent, D. P. (1991). Invariance and neural nets. *IEEE Transactions on Neural Networks*, *2*, 489–508.
- Becker, S. (1996). Models of cortical self-organisation. *Network: Comput. Neural Syst.*, *7*, 7–31.
- Becker, S. (1999). Implicit learning in 3D object recognition: The importance of temporal context. *Neur. Comput.*, *11*, 347–374.
- Becker, S., & Hinton, G. E. (1992). Self-organizing neural network that discovers surfaces in random-dot stereograms. *Nature*, *355*, 161–163.
- Bernander, O., Koch, C., & Douglas, R. J. (1994). Amplification and linearization of distal synaptic input to cortical pyramidal cells. *J. Neurophysiol.*, *72*, 2743–2753.
- Betsch, B. Y., Einhäuser, W., Körding, P. K., & König, P. (2001). *What a cat sees—statistics of natural videos*. Manuscript submitted for publication.
- Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford: Oxford University Press.
- Booth, M. C., & Rolls, E. T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cereb. Cortex*, *8*, 510–523.
- Bradschi G., & Grossberg, S. (1995). Fast-learning viewnet architectures for recognizing three-dimensional objects from multiple two-dimensional views. *Neural Networks*, *8*, 1053–1800.
- Braitenberg, V. (1978). Cortical architectonics: General and areal. In M. A. B. Brezler & H. Petsch (Eds.), *Architectonics of the cerebral cortex*. New York: Raven.

- Buzsaki, G., & Kandel, A. (1998). Somadendritic backpropagation of action potentials in cortical pyramidal cells of the awake rat. *J. Neurophysiol.*, *79*, 1587–1591.
- Cash, S., & Yuste, R. (1999). Linear summation of excitatory inputs by CA1 pyramidal neurons. *Neuron*, *22*, 383–394.
- Cauler, L. J., & Connors, B. W. (1994). Synaptic physiology of horizontal afferents to layer I in slices of rat SI neocortex. *J. Neurosci.*, *14*, 751–762.
- de Sa, V. R. & Ballard, D. H. (1998) Category learning through multimodality sensing. *Neur. Comp.*, *10*, 1097–1117.
- Dudek, S. M., & Bear, M. F. (1991). Homosynaptic long term depression in area CA1 of hippocampus and the effects on NMDA receptor blockade. *Proc. Natl. Acad. Sci. U.S.A.*, *89*, 4363–4367.
- Eisele, M. (1997). Unsupervised learning of temporal constancies by pyramidal-type neurons. In S. W. Ellacon, J. C. Mason, & I. J. Anderson (Eds.), *Mathematics of neural networks* (pp. 171–175). Norwell, MA: Kluwer.
- Felleman, D., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex*, *1*, 1–47.
- Földiák, P. (1991). Learning invariance from transformation sequences. *Neural Comput.*, *3*, 194–200.
- Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.*, *36*, 193–202.
- Fukushima, K. (1988). Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural Networks*, *1*, 119–130.
- Fukushima, K. (1999). Self-organization of shift-invariant receptive fields. *Neural Networks*, *12*, 791–801.
- Graziano, M. S., & Gross, C. G. (1998). Spatial maps for the control of movement. *Curr. Opin. Neurobiol.*, *8*, 195–201.
- Hebb, D. O. (1949). *The organisation of behaviour*. New York: Wiley.
- Helmchen, F., Svoboda, K., Denk, W., & Tank, D. W. (1999). In vivo dendritic calcium dynamics in deep-layer cortical pyramidal neurons. *Nat. Neurosci.*, *2*, 989–996.
- Hinton, G. E. (1987). Learning translation invariant recognition in a massively parallel network. In G. Goos & J. Hartmanis (Eds.), *PARLE: Parallel architectures and languages Europe* (pp. 1–13) Berlin: Springer-Verlag.
- Hinton, G. E. (1989). Connectionist learning procedures. *Artificial Intelligence*, *40*, 185–234.
- Hirsch, J. A., Alonso, J. M., & Reid, R. C. (1995). Visually evoked calcium action potentials in cat striate cortex. *Nature*, *378*, 612–616.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, *160*, 106–154.
- Hubel, D. H., & Wiesel, T. N. (1998). Early exploration of the visual cortex. *Neuron*, *20*, 401–412.
- Hyvarinen, A., & Hoyer, P. O. (2000). Emergence of phase and shift invariant features by decomposition of natural images into independent feature subspaces. *Neural Comput.*, *12*, 1705–1720.

- Johnston D., Magee, J. C., Colbert, C. M., & Cristie, B. R. (1996). Active properties of neuronal dendrites. *Ann. Rev. Neurosci.*, *19*, 165–186.
- Kay, J., Floreano, D., & Phillips, W. A. (1998). Contextually guided unsupervised learning using local multivariate binary processors. *Neural Networks*, *11*, 117–140.
- Kim, U., Sanchez-Vives, M. V., & McCormick, D. A. (1997). Functional dynamics of GABAergic inhibition in the thalamus. *Science*, *278*, 130–134.
- Kobatake E., & Tanaka K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J. Neurophysiol.*, *71*, 856–867.
- König, P., Engel, A. K., Roelfsma, P. R., & Singer, W. (1995). How precise is neuronal synchronization? *Neural Comp.*, *7*, 469–485.
- Körding, K. P., & König, P. (2000a). Learning with two sites of synaptic integration. *Network: Comput. Neural Syst.*, *11*, 1–15.
- Körding, K. P., & König, P. (2000b). A learning rule for dynamic recruitment and decorrelation. *Neural Networks*, *13*, 1–9.
- Larkum, M. E., Kaiser, K. M., & Sakmann, B. (1999). Calcium electrogenesis in distal apical dendrites of layer 5 pyramidal cells at a critical frequency of backpropagating action potentials. *Proc. Natl. Acad. Sci. USA*, *96*, 14600–14604.
- Larkum, M. E., Zhu, J. J., & Sakmann, B. (1999). A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature*, *398*, 338–341.
- Le Cun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E, Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Comput.*, *1*, 541–551.
- Markram, H., Lübke, J., Frotscher, M., & Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, *275*, 213–215.
- Mel, B. W. (1993). Synaptic integration in an excitable dendritic tree. *J. Neurophysiol.*, *70*, 1086–1101.
- Mel, B. W., Ruderman, D. L., & Archie, K. A. (1998). Translation-invariant tuning in visual “complex” cells could derive from intradendritic computations. *J. Neurosci.*, *18*, 4325–4334.
- O’Reilly, R. C., & Johnson, M. H. (1994). Object recognition and sensitive periods: A computational analysis of visual imprinting. *Neur. Comp.*, *6*, 357–389.
- Perrett, D. I., Oram, M. W., Harries, M. H., Bevan, R., Hietanen, J. K., Venson, P. J., & Thomas, S. (1991). Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Exp. Brain Res.*, *86*, 159–173.
- Phillips, W. A., Kay, J., & Smyth, D. (1995). The discovery of structure by multi-stream networks of local processors with contextual guidance. *Network: Comput. Neural Syst.*, *6*, 225–246.
- Pike, F. G., Meredith, R. M., Olding, A. W. A., & Paulsen O. (1999). Postsynaptic bursting is essential for “Hebbian” induction of associative long-term potentiation at excitatory synapses in rat hippocampus. *J. Phys.*, *518*, 571–576.
- Rao, R., & Sejnowski, T. J. (2000). Predictive sequence learning in recurrent neocortical circuits. In S. A. Solla, T. K. Leen, & K.-R. Müller (Eds.).

- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat. Neurosci.*, *2*, 1019–1025.
- Rolls, E. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Phil. Trans. R. Soc. Lond. B*, *335*, 11–21.
- Rolls, E., & Treves, A. (1997). *Neural networks and brain function*. Oxford: Oxford University Press.
- Salin, P. A., & Bullier, J. (1995). Corticocortical connections in the visual system: Structure and function. *Physiol. Rev.*, *75*, 107–154.
- Schiller, J., Schiller, Y., Stuart, G., & Sakmann, B. (1997). Calcium action potentials restricted to distal apical dendrites of rat neocortical pyramidal neurons. *J. Physiol. (Lond.)*, *505*, 605–616.
- Segev, I., & Rall, W. (1998) Excitable dendrites and spines: Earlier theoretical insights elucidate recent direct observations. *Trends Neurosci.*, *21*, 453–460.
- Siegel, M., Kording, K. P., & König, P. (2000). Integrating bottom-up and top-down sensory processing by somato-dendritic interactions. *J. Comput. Neurosci.*, *8*, 161–173.
- Softky, W. (1994). Sub-millisecond coincidence detection in active dendritic trees. *Neuroscience*, *58*, 13–41.
- Stone, J. V., & Bray A. J. (1995). A learning rule for extracting spatio-temporal invariances. *Network: Comput. Neural Syst.*, *6*, 429–436.
- Stuart, G. J., & Sakmann, B. (1994). Active propagation of somatic action potentials into neocortical pyramidal cell dendrites. *Nature*, *367*, 69–72.
- Stuart, G. J., Schiller, J., & Sakmann, B. (1997). Action potential initiation and propagation in rat neocortical pyramidal neurons. *J. Physiol. (Lond.)*, *505*, 617–632.
- Stuart, G. J., & Spruston, N. (1998). Determinants of voltage attenuation in neocortical pyramidal neuron dendrites. *J. Neurosci.*, *18*, 3501–3510.
- Wang, G., Tanaka, K., & Tanifuji, M. (1996). Optical imaging of functional organization in the monkey inferotemporal cortex. *Science*, *272*, 1665–1668.
- Williams, S. R., & Stuart G. J. (1999). Mechanisms and consequences of action potential burst firing in rat neocortical pyramidal neurons. *J. Physiol. (Lond.)*, *521*, 467–482.
- Zeki, S., & Shipp, S. (1988). The functional logic of cortical connections. *Nature*, *335*, 311–317.