

Synchronization of Neuronal Assemblies in Reciprocally Connected Cortical Areas ¹

Andreas Knoblauch and Günther Palm

Department of Neural Information Processing

University of Ulm, D-89069 Ulm, Germany

Tel: (0049)-731-50-24151; FAX: (0049)-731-50-24156

e-mail: (knoblauch,palm)@neuro.informatik.uni-ulm.de

Summary

To investigate scene segmentation in the visual system we present a model of two reciprocally connected visual areas comprising spiking neurons. The peripheral area P is modeled similar to the primary visual cortex, while the central area C is modeled as an associative memory representing stimulus objects according to Hebbian learning. Without feedback from area C, spikes corresponding to stimulus representations in P are synchronized only locally (slow state). Feedback from C can induce fast oscillations and an increase of synchronization ranges (fast state). Presenting a superposition of several stimulus objects, scene segmentation happens on a time scale of hundreds of milliseconds by alternating epochs of the slow and fast state, where neurons representing the same object are simultaneously in the fast state. We relate our simulation results to various phenomena observed in neurophysiological experiments, such as stimulus-dependent synchronization of fast oscillations, synchronization on different time scales, ongoing activity, and attention-dependent neural activity.

Key words: *Binding problem, temporal correlation hypothesis, spike synchronization, gamma oscillations, scene segmentation*

1 Introduction

One of the most fundamental and central questions in systemic neuroscience and brain research is the question of representation of objects or thoughts in the cortex of sufficiently complex animals like monkeys or men. Probably the most reasonable idea on this issue goes back to the psychologist Donald Hebb. Concepts or thoughts are represented in terms of the coincident activation of groups of neurons called *cell-assemblies* (Hebb, 1949). One immediate argument against this idea which has been put forward many times is the *superposition problem*: if two or more assemblies are activated at the same time, how can they be segmented into the individual assemblies? The answer to this problem is provided by the strong mutual excitation between the neurons belonging to

¹appeared in: Theory in Biosciences (2003) 122:37–54; (c) Urban&Fischer Verlag

the same assembly, probably acquired by Hebbian learning or auto-association. This has been worked out in several theoretical papers (Willshaw et al., 1969; Braitenberg, 1978; Palm, 1980, 1982, 1990; Sommer and Palm, 1999) which essentially show that by controlling the total activity within a certain region of the cortex, for example by unspecific inhibition, the superposition problem, i.e. the activation of two or more assemblies at the same time, can be avoided. Furthermore this can be achieved even when a very large number of assemblies (sparse activation patterns) are stored by auto-association.

But is this really a solution? May it not be necessary or at least useful in some situations to activate two or more assemblies at the same time in the same local cortical area and not to confuse them? In a visual scene, for example, there may be several objects present close to each other at the same time: If each of them is represented by an assembly that contains a subassembly in a visual area, and if the whole scene is also represented by an assembly, then these visual sub-assemblies should all be part of the “scene assembly” and therefore be activated at the same time in the same local region of visual cortex. This problem becomes more acute in higher visual areas where the assemblies corresponding to the different objects cannot be assumed to be spatially separated in the cortex due to a topographical visual representation with a sufficiently large magnification factor as in V1 or V2. But even in V1 or V2 neighboring neurons may be activated by different objects, where they are close together or even partially occluding each other. This particular version of the superposition problem has suggested a closer look on the timescale on which coincidence of activity is interpreted. It could be that there is a succession of coincident activity moving from one assembly to the next within a few milliseconds (i.e. on the timescale of single spikes), and also coincident activity of the whole set on the timescale of 100msec or of increased firing rates. This is the strong version of the temporal correlation hypothesis (TCH) (von der Malsburg, 1986) which was put forward as a solution to the superposition problem (also called the ‘binding problem’).

Actually, even a longer time window of up to three seconds (Pöppel, 1994) may be involved, for example, when we look at a complex picture, or obtain an impression of the room we have just entered, which may require several fixations and involve long-term memorized knowledge. This is a quite common situation, where from the point of view of assembly theory (Braitenberg, 1978; Palm, 1982) it may be reasonable to assume that a global assembly representing the situation as a whole is activated in the cortex which contains several subassemblies representing several individual objects. These subassemblies are probably only partially activated and in particular it is unlikely that several of them are activated at the same time (narrow time-window) in the same place (local cortical area, within 1-2mm of cortex), because that would violate the constraints on the total local activation. Using threshold control (or eye-movements) we can focus our attention (or our gaze) on each of these objects in turn and activate the corresponding assembly more completely, thereby reducing the activation of other assemblies representing other objects that are part of the same global situation.

The TCH was experimentally supported by the observation of high-frequency

gamma (40-60 Hz) oscillations in the visual cortex apparently reflecting global stimulus properties (Eckhorn, 1994; Singer and Gray, 1995). Evidence for this oscillatory activity was found in EEG or local field potential recordings and also in single-unit auto- or cross-correlograms. While the general idea of interpreting coincidence in assemblies on the timescale of milliseconds or spikes has been pursued by many if not most researchers elaborating the Hebbian assembly concept in the temporal domain (Abeles, 1982; von der Malsburg, 1986; Palm, 1990; Ritz et al., 1994; Aertsen et al., 1994; Wennekers and Palm, 2000), the special use of fine-timing made in the context of the visual 'binding problem', also related to the idea of 'phase-coding' (activating different assemblies at different phases of a single underlying oscillation), has been subject to considerable controversies.

The main purpose of this paper is to show that the original Hebbian ideas and the corresponding models based on associative memory almost automatically and naturally solve the "binding problem" by temporal correlation, when they are implemented in biologically realistic networks of spiking neurons. Global cortical assemblies are formed automatically when one assumes Hebbian synapses both in the long-range cortico-cortical connections between different areas and in the excitatory short-range connections within areas. Since most of the cortico-cortical connections appear to be bidirectional, a good starting point is to analyse two bidirectionally connected local areas. Here we essentially interpret our two cortical areas as two local spots of two visual areas, a more peripheral and a more central one, with a high and low spatial magnification factor, respectively. For our neuron-model, we use a spike generation mechanism with a dynamic threshold quite similar to common leaky integrate-and-fire models. The corresponding simulations and a few more extended simulations and extrapolations are summarized in this contribution.

The paper is organized as follows: In section 2 we describe our model in detail. Sections 3 and 4 show some results of our simulations and relate them to experimental observations. In the conclusions we discuss the resulting picture of information processing in the cortex in terms of cell assemblies.

2 Modeling

For simulations we chose a one-point spiking neuron model including refractoriness and habituation. The network consists of three areas (R, P and C) each composed of several neuron populations (Figure 1). Details of the neuron and network model are described in (Knoblauch and Palm, 2001, 2002a,b).

FIGURE 1 ABOUT HERE

In the retinal area R input patterns corresponding to stimulus objects in the visual field are represented in a 100×100 bitmap. The two cortical areas P (primary visual cortex) and C (central associative area) are both modelled with one excitatory (P and C) and two inhibitory neuron populations (P^S , P^T , C^S , and C^T), where only one of them receives extra-areal input (P^S and C^S). Only the excitatory connections are specific (with respect to orientations in area P, and with respect to the pattern representations in area C) according

to Hebbian learning. This architecture was motivated by the requirement of an efficient threshold control, especially for the associative area C (Knoblauch and Palm, 2001), and is also biologically plausible (Braitenberg and Schüz, 1991).

Each neuron population in area P has size 100×100 . Connections from R and inside P are modeled corresponding to the subsystem of orientation selective columns in the primary visual cortex. The internal excitatory connections from P to populations P and P^S couple specifically neurons which have similar orientation preferences and additionally are near neighbored or collinearly aligned. This is as expected from Hebbian correlation learning during the presence of stimuli rich of contours, and results in a patchy representation of stimulus objects (Fig. 1a): At a certain location in P, only the neurons with orientation preferences best matching the pattern in R get strongly activated, while neighboring cells are suppressed by recurrent inhibition. To prevent artificial long-range correlations we have also incorporated in our model temporal and spatial correlated noise as ubiquitously found in experiments (e.g. Tsodyks et al. (1999)). Due to computational limitations we replaced the two inhibitory populations P^S and P^T in some of the simulations by one gradual inhibitory population.

Area C (central visual area) is modeled as a fully connected auto-associative network according to Hebbian learning of topographic random representations of stimulus objects. As can be seen from the triangle representation in area C (Fig. 1a), the original stimulus shape is only rawly preserved. Populations C and C^S have size 40×40 , while population C^T has only size 20×20 (see Knoblauch and Palm (2001) for details). Areas P and C are reciprocally connected by a topographic hetero-association according to Hebbian learning of corresponding stimulus representations in P and C.

We use simple shapes as test stimuli, e.g. a triangle, a rectangle or an ellipse. In different simulations we present single stimuli to focus on the general behavior of the network (e.g. appearance of oscillations and distance dependent synchronization), or superpositions of several stimuli to investigate competition and segmentation properties of the model. To investigate more global model properties we simulated also a larger variant of the model where areas P and C have sizes 30×500 and 15×250 , and where also area C is organized topographically (Knoblauch and Palm, 2002b).

3 Basic simulation results

3.1 Single Stimulus

To investigate the general behavior of area P, we first simulated the network *without any feedback* from C: Activity in both areas then is oscillatory due to the interaction of excitatory and inhibitory pools of neurons. In the primary area P one obtains activated patches along the contour of the stimulus shape (cf. Fig.1a). Correlation analysis of multi-unit spike activity from different such patches reveals that only activity from neighbored patches is synchronized (modulated correlograms with central peaks) while activity from distant patches

is uncorrelated (flat correlograms, cf. Fig. 2a,c).

FIGURE 2 ABOUT HERE

The situation changes if the *complete model* is simulated including the feedback from area C to area P. Then also correlograms for activity from distant patches show significant central peaks indicating an enlarged synchronization range (Fig. 2a,c).

In summary the simulations show that spike synchronization (on a time scale of milliseconds) under realistic noise conditions is only possible for neurons in either directly (neighbored in area P) or indirectly (by divergent feedback from area C) connected cortical regions. This is consistent with neurophysiological results where synchronization seems to be restricted to synaptically connected regions (some millimeters within visual cortex, cf. Eckhorn (1994)), but it obviously contradicts the strong version of the TCH: This states that global assemblies of neurons should be synchronized on the time scale of milliseconds despite the lack of direct connections between parts of the assembly. However, as we will see later, global synchronization is indeed possible on a larger time scale of tens to hundreds of milliseconds.

3.2 Multiple stimuli

According to the TCH synchronous fast (γ) oscillations were suggested to solve the binding problem. In this subsection we test our model using a superposition of three test stimuli, a triangle, an ellipse, and a rectangle. Here the representation of each stimulus in P is a certain set of orientation selective patches (cf. Fig. 3c), and the binding problem is to coordinate activity of patches that belong to one representation in P as well as corresponding representations in P and C. Previous models solved this kind of binding problem for example by phase coding (e.g. Ritz et al. (1994)). That means, assemblies representing different stimuli are activated during different phases of *one* γ period. However, there exists little experimental evidence for this kind of coding. Rather, experiments investigating the representations of several simultaneously presented stimuli indicate *uncorrelated* activity between assemblies for the different stimuli corresponding with flat correlograms (Singer and Gray, 1995). In case of a phase-coding mechanism the correlograms should be modulated with certain shifts of the central peaks.

FIGURE 3 ABOUT HERE

Figure 3ab shows simultaneous recordings of spike activity from neurons in P and C. Multi-unit spike activity is depicted separately for the different (sub)assemblies representing the three stimuli (triangle, ellipse and rectangle). The recordings of one assembly in C (Fig. 3a, e.g. the bottom row for the rectangle assembly) show periods of fast oscillatory activity lasting for a few hundred milliseconds alternating with longer periods of essential silence. Comparing the three recordings from C, one observes that only one assembly is in the fast state at a time. This results from the structure of the associative memory in area C, which prefers separated activation of the three addressed assemblies (Knoblauch

and Palm, 2001).

Multi-unit activity of patches belonging to the three assemblies in area P reveals similar results (Fig. 3b) as in area C. The recordings show periods of fast and precise (little phase jitter) oscillatory activity (*fast state*) lasting again for a few hundred milliseconds alternating with periods of relatively slow and unordered activity (*slow state*). When looking at different patches of one assembly one observes that activity is synchronized only during the fast state, but asynchronous during the slow state. Moreover, corresponding assemblies in P and C are simultaneously either in the fast or in the slow state. Comparing with the single stimulus conditions, one finds that the slow state is very similar to the results without feedback from the higher area, and that the fast state is very similar to the results of the complete model (Knoblauch and Palm, 2002a).

Mean firing rates show a superposition of the assemblies representing the three stimuli (Fig. 3c,d). Correlation analysis reveals that spike activity within one assembly is synchronized on a timescale of milliseconds (modulated correlograms with central peaks, see Fig. 3e for the triangle assembly in area P), while spikes of different assemblies are uncorrelated (flat correlograms, see Fig. 3f for the cross correlation of triangle and ellipse assemblies in P). This is consistent with findings of specifically modulated and unmodulated correlograms in the experiments described above using multiple stimulus objects (see section 4.1).

Cross correlations for activities of corresponding assemblies in P and C show modulations with the central peak shifted for about half a period (Fig. 3g for the triangle assembly), which occurs similarly for a single stimulus in the complete model (data not shown, cf. Knoblauch and Palm (2002b)). However, fast oscillatory activity in different visual areas is consistently reported to be synchronized with *zero* phase lag. In section 4.2 we will discuss this apparent inconsistency in more detail.

Figure 3h shows the cross correlations for two distant triangle patches in area P (patches as in Fig. 2a,b), but on a larger time scale. While the correlations on the finer time scale looked very similar to Fig. 2b, here the correlogram shows one broad central peak indicating non-oscillatory synchronization on a larger time scale. In our simulations these broad peaks result from the switching between two states as described above, and they occur centered on the time origin even for activity from different areas. In section 4.3 these results are further discussed and related to findings of broad peaks in physiological data.

4 Relation to neurophysiological experiments

4.1 Stimulus-specific correlated fast oscillatory activity

In section 3.2 we investigated scene segmentation using a superposition of three different stimulus objects (Fig. 3). The neuron groups representing one stimulus tended to be simultaneously in the so-called slow state or the fast state, thereby segmenting the superposition. Cross correlations between cell groups coding features corresponding to a common stimulus were modulated, while

cross correlations remained flat, if the cell groups coded features corresponding to different stimuli. These results are consistent with experimental findings of specifically modulated and unmodulated correlograms (Singer and Gray, 1995). In (Knoblauch and Palm, 2002a) we simulated a stimulus configuration in analogy to these experiments (using static instead of moving stimuli). We recorded simultaneously from two locations testing two stimulus configurations. In the first configuration the two locations were stimulated by one object, while in the second configuration they were stimulated by two different objects. This is reflected in modulated correlograms for the first configuration and flat correlograms for the second.

In earlier proposed *phase coding models* (e.g. Ritz et al. (1994)) simultaneously presented stimuli are segmented by serially activating the corresponding assemblies in a fixed order within one gamma-period (about 30 msec). However, these models inconsistently would predict correlograms with shifted peaks if the two groups code different entities.

4.2 Interareal synchronization of fast activity

Interpreting area C for example as V2 or V4, the finding of anti-phase interactions for the fast oscillatory synchronization between areas P and C (Fig. 3g) is inconsistent with experimental results of zero-phase synchronization between connected visual areas (e.g. Frien et al. (1994)).

Further simulations (Knoblauch and Palm (2002b), cf. Ritz et al. (1994)) showed that activity from the two areas is correlated with either zero-phase or anti-phase depending on the inter-areal transmission delays and the oscillation period T . Small delays support zero-phase, medium delays around $T/2$ result in anti-phase, whereas large delays around T support again zero-phase.

For the simulations so far we used transmission delays around 7msec (2.5msec - 10msec, see Knoblauch and Palm (2002b) for details), which is about $T/2$ for an oscillation frequency around 60Hz. Using smaller inter-areal delays yields zero-phase correlations as in the neuro-physiological experiments, but then our model does not exhibit the scene segmentation capability as described before, because inhibition cannot be bridged by the small delays. Since experiments rather support broad delay distributions reaching values in the tens of milliseconds (Swadlow, 2000) we simulated a model variant with a bimodal distribution of inter-areal transmission delays (including very small delays corresponding to fastly conducting myelinated axons but also broadly distributed large delays corresponding to slowly conducting unmyelinated axons). We obtained both (near) zero-phase correlations *and* scene segmentation (Knoblauch and Palm, 2002b). Moreover, we proved in a further simulation study (Knoblauch and Sommer, 2003) that spike-time-dependent synaptic plasticity (Markram et al., 1997) in an oscillatory regime strengthens synapses with delays in the range of one oscillation period such that zero-phase-lag is stabilized even for realistic delay distributions which would otherwise lead to anti-phase correlations.

4.3 Synchronization on a larger time scale

Correlation analyses of spike activity from neurophysiological recordings reveal evidence for synchronization on different time scales (Nelson et al., 1992). Widths of correlogram peaks show a tripartite distribution classified as T(ower) (width at half height $< 16\text{ms}$), C(astle) ($16\text{--}180\text{ms}$) and H(ill) peaks ($> 180\text{ms}$). Both T and C peaks seem to depend on cortico-cortical connections and thus the broad peaks in our model (Fig.3h) may be related to the C-peaks.

In our model the widths of the broad peaks depend on the duration of the fast state of oscillatory activity (section 3.2), and the switching between the slow and fast state is obviously due to habituation of activated neurons. In simulations (Knoblauch and Palm, 2002b) we manipulated the duration of the fast state by varying the habituation parameters of the neurons. For strong habituation we obtained durations down to 20-50msec where the enhanced periods often degenerated to a single retrieval cycle. For weak habituation the duration of the fast state periods increased up to seconds, and could result in the sustained activation of a single assembly without switching between different assemblies.

The correlograms found in neurophysiological experiments were either flat or showed central peaks, while oscillatory modulations or central troughs were observed rarely. According to our model one should expect oscillatory modulations with central troughs at least for two competing stimuli. To investigate under which conditions our model reproduces the experimentally observed correlograms we used up to seven overlapping stimuli (Knoblauch and Palm, 2002b). While broad peaks were absent for single stimuli, and two stimuli induced anti-phasic correlations, we obtained the experimental correlogram peaks if many assemblies are activated at the same time. Then correlograms remain flat for different assemblies and show single broad peaks within an assembly.

A possible conclusion could be that the brain is actually in a 'many-activated-assemblies' regime even without stimulation, where ongoing activity supposedly wanders permanently through different local assemblies (Tsodyks et al., 1999). Indeed, the behavior of our model without stimulation but with enhanced noise level is very similar to the case of many simultaneously presented stimuli. In this view, a stimulus causes the corresponding assembly to spend more time in the fast state as in the other assemblies. This would be in line with recent findings of stimulus dependent two-state fluctuations of membrane potentials (Anderson et al., 2000) and could also explain neurophysiological results described next.

4.4 Attention and biased competition

The results from the previous sections suggest an involvement of attentional processes in scene segmentation. The prevalence of one of the three stimulus objects was switched on a time scale of tens to hundreds of milliseconds (fast and slow state, see section 3.2). This may be interpreted as self-generated attention switching serially from one object to the next. For the following simulations we modeled the top-down attentional bias explicitly as additional tonic excitation to the neurons of one selected assembly in area C (Treisman, 1998). We tried

to simulate a scenario as in experiments described in (Reynolds and Desimone, 1999). In this scenario two stimuli in the receptive field of a single neuron were used, a preferred and a poor stimulus. Presenting only the preferred stimulus, the neuron exhibits a strong response, while the response is weak, if only the poor stimulus is used. Interestingly, for the superposition of the two stimuli, the response lies between the responses for single stimuli, indicating competition between the two stimuli. If attention is directed to one of the stimuli, the neuron responds as if only the attended stimulus were present. Thus the effect of attention is similar to a filter that eliminates or weakens unattended stimuli. Another result of the neurophysiological experiments was that the described effect scales with the receptive field size, i.e. the effect was stronger in higher visual areas (e.g. V4 or IT), and weaker in lower visual areas (e.g. V1 or V2).

In (Knoblauch and Palm, 2002b) we show that our model can reproduce experimentally measured neuron responses quite well. It explains the competition and the filter property by two activation states (fast and slow states). For one stimulus a neuron is in the fast state all the time, while with two competing stimuli the neuron is only about half the time in the fast state; attentional bias, however, brings the neuron into the fast state again.

4.5 Synchronization on a larger spatial scale

The model as proposed so far should be interpreted as the bidirectional connection of two small patches of cortex. In the real visual system many such structures may be arranged in parallel. To account for more global interactions we simulated a model variant with larger areas (area P size 30×500 ; area C size 15×250), where area C is modeled as a *topographic* associative memory (Knoblauch and Palm, 2002b). While the structure of area P remained unchanged, the connection kernels within area C and between areas P and C were reduced to sizes 31×21 and 15×11 , respectively. The kernel sizes in y-direction were larger to model competition in y-direction and topography in x-direction. As stimuli we used three long bars that extended over the whole areas and were represented by three learned assemblies in area C. Fig. 4 shows results from a single simulation of the model stimulated with all three bars simultaneously. At time $t = 4\text{sec}$ attention was directed for three seconds to the upper bar, while attention was absent in the rest of the simulation.

When attention is directed to one of the three bars the whole corresponding assembly enters the fast state, similarly as if only the attended bar were present in the visual field (cf. section 4.4). In addition, the globally synchronized oscillations of the smaller model (see section 3.2) mutate to waves of activity moving fastly along the bar. At a given time the oscillation phase is a continuous function of the location along the bar representation (Fig.4b) reflecting the global shape of the stimulus. However, fast activity in P and C is synchronized only locally corresponding to the feedback range of intra-areal and inter-areal connections, and correlation analysis reveals that the central peaks of (long-term) correlograms decay with distance (Knoblauch and Palm, 2002b). Moreover, short-term correlograms (time window 50msec) show that the central peaks of

the long-term correlograms result mainly from averaging over phase-shifts that increase with distance from the recording sites. A similar effect can also be observed for the broad peaks (section 4.3).

Without attention the representations of the three bars compete in the higher area C. At a fixed location the situation is very similar to the local model (section 3.2): The assembly corresponding to the most salient stimulus wins and enters the fast state. After some time (e.g. some hundred milliseconds) the assembly habituates and returns to the slow state, such that another assembly can enter the fast state. However, for the global model it turns out that the assembly switching is a rather local property. At different locations different assemblies can be in the fast state at the same time.

Correlation analysis reveals that synchronization on the larger time scale extends further (Knoblauch and Palm, 2002b) than synchronization on the finer time scale. Actually it turns out that the synchronization range depends mainly on the duration of the fast state, and is therefore independent of the extension of the underlying synaptic connections. It could therefore be manipulated, e.g., by modulatory synaptic input or attention. This long-range synchronization on a larger scale of space and time might be an important property of the visual system to integrate local features of possibly new objects in the visual field. Although synaptic connectivity is local, global synchronization can be achieved as soon as neighboring local stimulus (sub)representations (e.g., in a topographical area) are compatible with each other.

5 Conclusions

To investigate synchronization of distributed neural assemblies we have proposed a model of two interacting cortical areas which is compatible with several lines of experimentally observed phenomena such as synchronization on a fast time scale (Eckhorn, 1994; Singer and Gray, 1995) or a slow time scale (Nelson et al., 1992), ongoing activity in optical recordings (Tsodyks et al., 1999), two-state fluctuations of membrane potentials (Anderson et al., 2000), and attentional or biased competition (Reynolds and Desimone, 1999).

Neural assemblies are distributed representations that actually need binding of their components, probably by synchronization of spike activity. Before we discuss possible solutions to the binding problem (BP) in the context of our simulations, note that the BP occurs at different levels of difficulty.

1. Binding between neurons representing externally correlated or co-occurring features.
2. Binding across different, often unrelated feature dimensions (such as form and color) within one assembly.
3. Dynamic binding (e.g. role assignment) relating different assemblies, where possibly fast synaptic plasticity is required.

The simplest BP1 is to bind together corresponding subrepresentations like the features of a complex object and perhaps also an abstract representation of the whole object in a higher area. This problem is relatively simple because we can assume that specific synaptic connections exist between the subrepresentations. So far, our model accounts only for this simplest BP: The features are the orientation patches in the peripheral area P, and the abstract representation are the object assemblies in the central area C. Another instructive example is the famous picture of the dalmation dog composed of a chaos of black patches, cf. Eckhorn (1994). Here one can observe that BP1 occurs in two subforms: It is rather difficult if the abstract representation does not (yet) exist as it is difficult to see the dog for the first time. In contrast, if recognized only once, the dog is much easier to recognize the next time. So we can assume that actually an abstract representation of this picture was allocated in our brains that facilitates the binding of the black patches. However, we are also capable of recognizing the dalmation dog if we see the picture for the first time, although it will take some time. We believe that binding in our larger model (Fig.4) — where no global representations of the bars exist — is similar to this process (although much easier). A refined coding strategy would not require an extra code for each object. For example different objects could be segmented by local properties alone, such as continuously aligned contours, cf., e.g. Neumann and Sepp (1999). A process like the slow waves in our simulations could help to segment a visual scene, although on a relatively large time scale.

A more difficult problem is BP2 that occurs if binding across different feature dimensions (such as form and color, cf. Fig. 5) is required. In this case we cannot always expect specific synaptic connections between the different processing streams. For example artificial objects in the visual field can occur with arbitrary colors, and combinatory coding each of n forms and each of m colors would require $n \times m$ representations. Separating the feature dimensions, however, requires only $n+m$ representations plus an adequate binding mechanism. Thus it would be very inefficient and therefore surprising if the brain would not take advantage of the possibility of separating different feature dimensions.

Figure 5 sketches a more global model to solve at least BP1 and BP2 that follows from our simulations. Binding between subassemblies of two connected areas, i.e. BP1, is accomplished by reciprocal excitation leading to the fast synchronized activation state. Local assemblies in one area are competing with each other and only the winning assembly is routed to the next higher processing stage. The range of competition is larger for higher areas. The bias of external excitation (e.g. attention) can select the subassemblies to be processed globally. While the assembly corresponding to a “non-winning” (non-attended) object is suppressed in a high area, pre-attentive processing may still be possible in a lower area where the corresponding assemblies do not compete due to spatial separation. Binding across different processing streams may be possible due to the indirect reciprocal connections via the primary areas and/or bottom-up synchronization mechanisms, e.g., locally correlated random fluctuations (Horn et al., 1991), or a localized ‘searchlight’ (Crick and Koch, 1990).

An even more serious problem is BP3 which occurs if dynamic binding is

required between different entities. This occurs for example if one tries to understand the sentence “Tom beats Fritz”. Here one can assume the existence of previously learned representations of the involved persons (Tom and Fritz), and also representations of the meaning of “to beat” which involves representations of a beater and a victim, but no prejudicial bias of the role of Tom and Fritz. Here the binding problem is to assign these roles dynamically to Tom and Fritz. One solution to BP3 could be short-term plasticity that would temporarily connect the representations ‘Tom’ with ‘beater’, and similarly ‘Fritz’ with ‘victim’. Although this fast synaptic plasticity (on a time scale below a second) was proposed already long ago (von der Malsburg, 1986) it is experimentally not easy to measure and its existence is still uncertain today. So the question arises if and how BP3 can be solved without fast synaptic plasticity.

Note that the structure of BP3 is very similar to BP2. BP2 involves for example understanding a visual scene containing a red apple and a green pear where temporary associations between ‘red’ and ‘apple’ and between ‘green’ and ‘pear’ are involved similarly as for BP3. As we discussed above BP2 can be solved by a topographical representation where basic feature representations of the apple and the pear are still connected with the correct color representations due to spatial separation. This spatial separation principle could also work for BP3 which would reveal it as a pseudo-problem. A possible solution is the following: “Tom beats Fritz” must be translated into a *topographical generic feature representation* where a generic feature representation of ‘Tom’ is connected with a generic feature representation of ‘beater’, and the same for ‘Fritz’ and ‘victim’. This could happen for example by visual imagination of a visual scene where Tom beats Fritz. However, since we seem to understand “Tom beats Fritz” even without visual imagination there may be involved also non-visual topographical representations. Although it is true that no synaptic plasticity is required by this solution for BP3 there is still a need for a kind of “working memory” to preserve e.g. an imagined scene by sustained activity.

Another solution to BP3 not requiring sustained activity or fast synaptic plasticity suggested by our simulations would rely on common habituation levels of corresponding representations. For example hearing “Tom beats ...” would first activate two assemblies ‘Tom’ and ‘beater’ at two different cortical locations for some time. Subsequently hearing “... beats Fritz” would similarly activate the representations ‘victim’ and ‘Fritz’ in the same two areas representing persons and roles, respectively. After suppressing activity in the two areas binding information would still be preserved in the habituation variables. If the two areas are unspecifically activated (assuming that all 4 assemblies are equally primed) one would obtain again ‘Tom’ and ‘beater’ because they are less habituated than ‘Fritz’ and ‘victim’.

Fast but decaying synaptic plasticity would improve binding in any case. How well the proposed binding mechanisms work remains to be seen in further simulations.

Acknowledgment

The authors want to thank Thomas Wennekers, Friedrich Sommer, and Friedemann Pulvermüller for valuable discussions. This work was partially supported by the Volkswagen foundation (VW-I/77 355), and the MirrorBot project of the European Union.

References

- Abeles, M. (1982). *Local cortical circuits*. Springer, Berlin Heidelberg New York.
- Aertsen, A., Erb, M., and Palm, G. (1994). Dynamics of functional coupling in the cerebral cortex: an attempt at a model-based interpretation. *Physica D*, 75:103–128.
- Anderson, J., Lampl, I., Reichova, I., Carandini, M., and Ferster, D. (2000). Stimulus dependence of two-state fluctuations of membrane potential in cat visual cortex. *Nature Neuroscience*, 3(6):617–621.
- Braitenberg, V. (1978). Cell assemblies in the cerebral cortex. In Heim, R. and Palm, G., editors, *Lecture notes in biomathematics (21). Theoretical approaches to complex systems.*, pages 171–188. Springer-Verlag, Berlin Heidelberg New York.
- Braitenberg, V. and Schüz, A. (1991). *Anatomy of the cortex. Statistics and geometry*. Springer-Verlag, Berlin.
- Crick, F. and Koch, C. (1990). Towards a neurobiological theory of consciousness. *Sem. Neurosci.*, 2:263–275.
- Eckhorn, R. (1994). Oscillatory and non-oscillatory synchronizations in the visual cortex and their possible roles in associations of visual features. In Pelt, J., Corner, M., Uylings, H., and Lopes da Silva, F., editors, *Progress in Brain Research*, volume 102, chapter 28, pages 405–426. Elsevier Science BV.
- Frien, A., Eckhorn, R., Bauer, R., Woelbern, T., and Kehr, H. (1994). Stimulus-specific fast oscillations at zero phase between visual areas V1 and V2 of awake monkey. *NeuroReport*, 5(17):2273–2277.
- Hebb, D. (1949). *The organization of behavior. A neuropsychological theory*. Wiley, New York.
- Horn, D., Sagi, D., and Usher, M. (1991). Segmentation, binding, and illusory conjunctions. *Neural Computation*, 3:510–525.
- Knoblauch, A. and Palm, G. (2001). Pattern separation and synchronization in spiking associative memories and visual areas. *Neural Networks*, 14:763–780.
- Knoblauch, A. and Palm, G. (2002a). Scene segmentation by spike synchronization in reciprocally connected visual areas. I. Local effects of cortical feedback. *Biological Cybernetics*, 87(3):151–167.
- Knoblauch, A. and Palm, G. (2002b). Scene segmentation by spike synchronization in reciprocally connected visual areas. II. Global assemblies and synchronization on larger space and time scales. *Biological Cybernetics*, 87(3):168–184.
- Knoblauch, A. and Sommer, F. (2003). Synaptic plasticity, conduction delays,

- and inter-area phase relations of spike activity in a model of reciprocally connected areas. *to appear in Neurocomputing*.
- Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, 275:213–215.
- Miller, R., editor (2000). *Time and the Brain*. Conceptual Advances in Brain research. Harwood Academic Publishers, Amsterdam.
- Nelson, J., Salin, P., Munk, M., Arzi, M., and Bullier, J. (1992). Spatial and temporal coherence in corticocortical connections: a cross-correlation study in areas 17 and 18 in the cat. *Visual Neuroscience*, 9:21–38.
- Neumann, H. and Sepp, W. (1999). Recurrent V1-V2 interaction in early visual boundary processing. *Biological Cybernetics*, 81:425–444.
- Palm, G. (1980). On associative memories. *Biological Cybernetics*, 36:19–31.
- Palm, G. (1982). *Neural Assemblies. An Alternative Approach to Artificial Intelligence*. Springer, Berlin.
- Palm, G. (1990). Cell assemblies as a guideline for brain research. *Concepts in Neuroscience*, 1:133–148.
- Pöppel, E. (1994). Temporal mechanisms in perception. *International Review of Neurobiology*, 37:185–202.
- Reynolds, J. and Desimone, R. (1999). The role of neural mechanisms of attention in solving the binding problem. *Neuron*, 24:19–29.
- Ritz, R., Gerstner, W., Fuentes, U., and van Hemmen, J. (1994). A biologically motivated and analytically soluble model of collective oscillations in the cortex. II. Applications to binding and pattern segmentation. *Biol. Cybern.*, 71:349–358.
- Singer, W. and Gray, C. (1995). Visual feature integration and the temporal correlation hypothesis. *Annu.Rev.Neurosci.*, 18:555–586.
- Sommer, F. and Palm, G. (1999). Improved bidirectional retrieval of sparse patterns stored by hebbian learning. *Neural Networks*, 12:281–297.
- Swadlow, H. (2000). Information flow along neocortical axons. In Miller (2000), chapter 4, pages 131–155.
- Treisman, A. (1998). Feature binding, attention and object perception. *Phil. Trans. R. Soc. London B*, 353:1295–1306.
- Tsodyks, M., Kenet, T., Grinvald, A., and Arieli, A. (1999). Linking spontaneous activity of single cortical neurons and the underlying functional architecture. *Science*, 286:1943–1946.
- von der Malsburg, C. (1986). Am I thinking assemblies? In Palm, G. and Aertsen, A., editors, *Brain Theory*, pages 161–176. Springer-Verlag, Berlin/Heidelberg.
- Wennekers, T. and Palm, G. (2000). Cell assemblies, associative memory and temporal structure in brain signals. In Miller (2000), chapter 10, pages 251–273.
- Willshaw, D., Buneman, O., and Longuet-Higgins, H. (1969). Non-holographic associative memory. *Nature*, 222:960–962.

Figure 1: Overview of the network model. **a**: Global connection scheme and the representations of a triangle stimulus in the different areas R, P, and C. Numbers in brackets correspond to the size of efferent kernels of individual neurons, e.g. a P-neuron projects to a patch in C with diameter 13 neurons. **b**: Individual neuron populations of the three areas and excitatory (arrows) and inhibitory connections between them.

Figure 2: Analysis of spike activity from area P when stimulating with a triangle (Fig.1a). Plots **a**, **c** correspond to the model with inactivated feedback from area C, while plots **b**, **d** correspond to the complete model. **a**, **b**: Cross correlograms for multi-unit activity from two (distant) activated patches in area P. **c**, **d**: Maximal correlations (maximized over time lag) vs patch distance for pairs of activated patches in P. Without feedback significant (central) peaks occur only for nearby patch pairs, while for the complete model also distant patch pairs are synchronized.

Figure 3: Analysis of spike activity when stimulating with a superposition of three stimuli (triangle, ellipse, rectangle; complete model). **a**: Summed spike activity in population C for the triangle (top row), ellipse (middle row) and rectangle assembly (bottom row). **b**: Summed spike records of 10 different activated patches representing the triangle (rows 6–9), the ellipse (rows 3–5) and the rectangle stimulus (rows 0–2) in population P. **c**, **d**: Spike rates in population P and C. **e**: Auto-correlogram (AC) for summed spike activity of the triangle assembly (all triangle patches) in population P. **f**: Cross-correlogram (CC) for summed spike activity from the triangle and ellipse assemblies. **g**: CC for the triangle (sub-)assemblies in populations P and C. **h**: CC for the same patch pair as in Fig.2a,b, but on a larger time scale.

Figure 4: Results for a simulation of the larger model stimulating with three bars simultaneously (simulation over 20sec). Between 4sec and 7sec after stimulus onset attention is directed to the (upper) bar 1. **a**: Spike rates in area P. 50 activated patches (black) reflect each of the bar stimuli. **b**: Plot of the membrane potentials vs time for the central neurons of the activated patches in area P for bar 1 (neurons at vertical position 5 and horizontal positions 8,28,...,498; the plot contains one line for each neuron). **c**: Plot of spike activity vs time for the central neurons of the activated patches for bar 1,2,3 (vertical positions 5,15,25 and horizontal positions 8,18,...,498). White patches indicate strong activation of neurons (fast state). Without attention the fast state is split up between the three bar representations due to competition (slow waves, fluctuating binding). Attention selects one of the bars (static binding).

Figure 5: Suggested binding by “slow waves” in a global model of three stages and two processing streams (form and color). Different colors are expressed by

different gray levels. Thick lines correspond to the fast enhanced activity state, thin lines to the slow state. In the topographically organized primary area objects separated in space do not compete with each other and are therefore in the fast state all the time. Competition, i.e., switching between different assemblies, occurs only at places where objects are overlapping. In the smaller topographically organized intermediate area nearby objects compete, and in the central area there is global competition. In the central area the “winning” (e.g. attended) assembly forces by feedback the corresponding assemblies in the peripheral areas to the fast state, and suppresses overlapping competing assemblies. Binding over different streams is accomplished indirectly via the primary area due to the reciprocal connections.

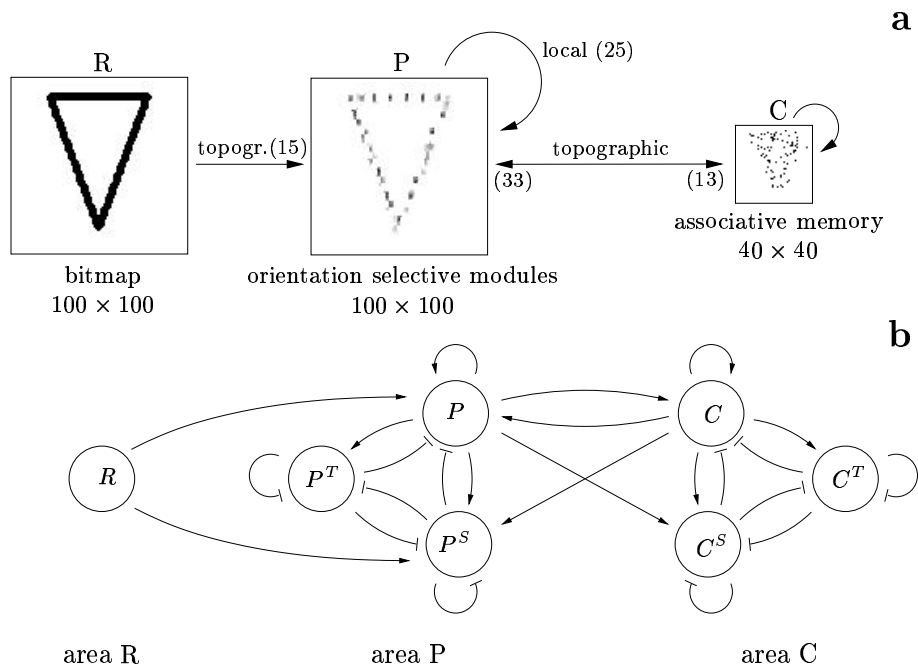


Figure 1:

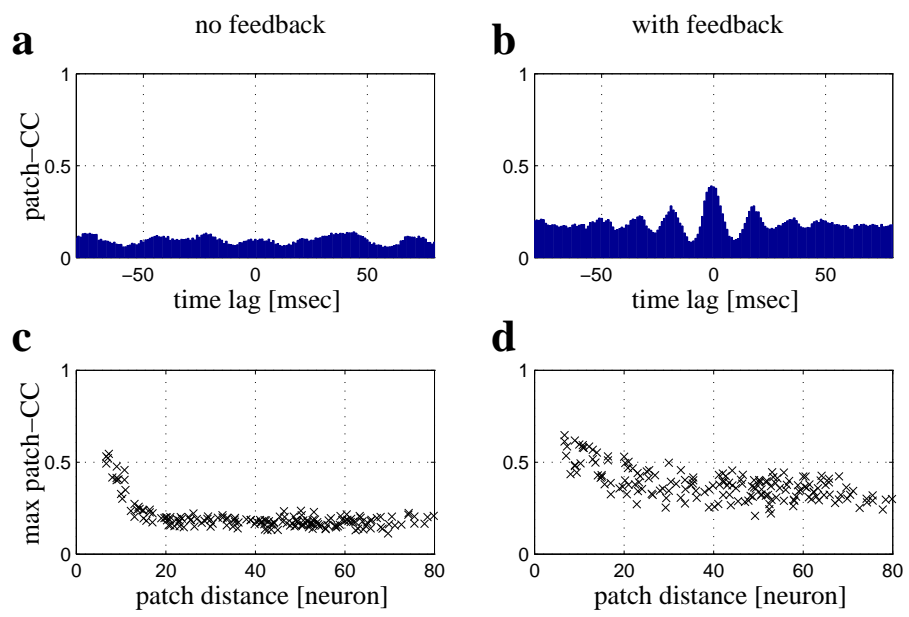


Figure 2:

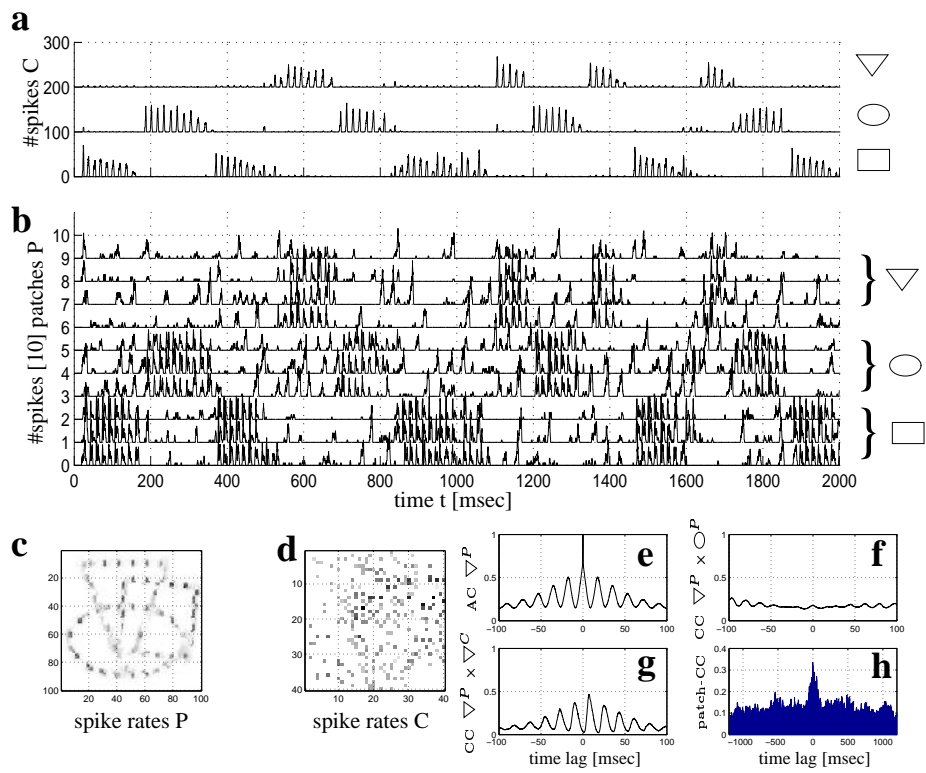


Figure 3:

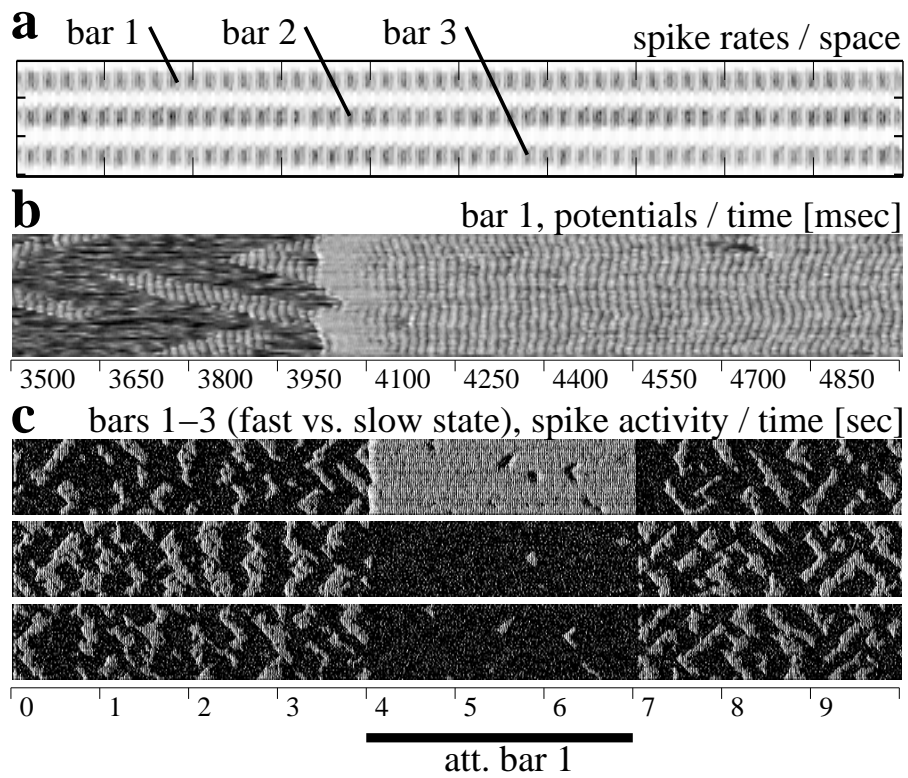


Figure 4:

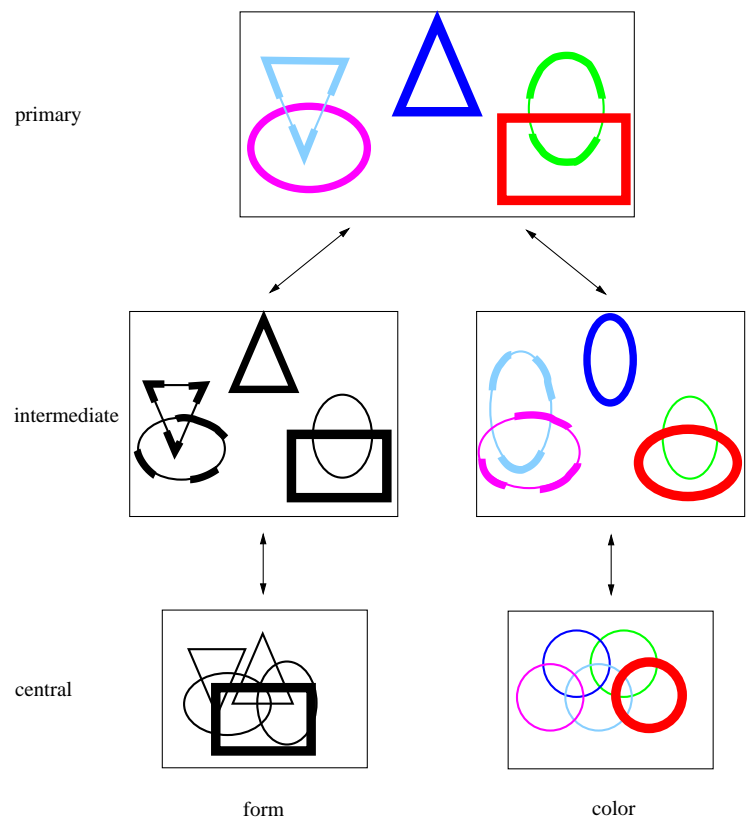


Figure 5: