Neural Computation 13(2), 357-387 (2001)

A Competitive Layer Model for Feature Binding and Sensory Segmentation

Heiko Wersing*, Jochen J. Steil, and Helge Ritter

University of Bielefeld, Faculty of Technology,
P.O.Box 100131, D-33501 Bielefeld, Germany

Abstract

We present a recurrent neural network for feature binding and sensory segmentation, the competitive layer model (CLM). The CLM uses topographically structured competitive and cooperative interactions in a layered network to partition a set of input features into salient groups. The dynamics is formulated within a standard additive recurrent network with linear threshold neurons. Contextual relations among features are coded by pairwise compatibilities which define an energy function to be minimized by the neural dynamics. Due to the usage of dynamical winner-take-all circuits the model gains more flexible response properties than spin models of segmentation by exploiting amplitude information in the grouping process. We prove analytic results on the convergence and stable attractors of the CLM, which generalize earlier results on winner-take-all networks, and incorporate deterministic annealing for robustness against local minima. The piecewise linear dynamics of the CLM allows a linear eigensubspace analysis which we use to analyze the dynamics of binding in conjunction with annealing. For the example of contour detection we show how the CLM can integrate figure-ground segmentation and grouping into a unified model.

1 Introduction

From the viewpoint of brain theory (von der Malsburg 1981; von der Malsburg 1995), feature binding may provide one of the basic sensory information processing principles. A key question is then what are the neural correlates of binding processes? In addition to the importance of this question regarding our understanding of brain function, there is also great interest in using similar mechanisms for pattern recognition applications like image segmentation and object recognition.

A large body of neural network research has focused on binding models based on temporally correlated neural activity (von der Malsburg 1981), stimulated by neurophysiological findings (Singer & Gray 1995; Eckhorn 1994) which support a functional role of synchronized activity in perceptual processes. Correlation-based feature binding has been modelled by phase-coupled oscillators (Baldi & Meir 1990; Sompolinsky, Golomb, & Kleinfeld 1991), where, however, problems are slow convergence and the necessity of all-to-all connections for robust synchrony. Other non-linear oscillator models have been developed (von der Malsburg & Buhmann 1992; Schillen & König 1994), but were only tested on

*new adress: HONDA R&D Europe (Germany), Future Technology Research, Carl-Legien-Str.30, 63073 Offenbach/Main, Germany, email: heiko.wersing@hre-ftr.f.rd.honda.co.jp

small networks and highly simplified test images. Relaxation oscillators (Somers & Kopell 1993; Terman & Wang 1995) show long range synchrony also with local connections and have been applied to region-based image segmentation (Wang & Terman 1997) and auditory segregation (Brown & Wang 1997). A problem is the limited number of groups which can be stably represented (about 5), which can only be overcome by introducing algorithmic abstractions of the original model. A phase averaging model with rule-based interactions has been applied to the extraction of contour saliency by Yen & Finkel (1998). Nevertheless, successful applications to the segmentation of real-world data are still rather exceptional (Wang & Terman 1997; Yen & Finkel 1998). This is mainly caused by the high dynamical complexity of these models, which makes their simulation costly and their analytic study a difficult task.

A different approach to feature binding are spin models, which have been developed for computer vision (Geman, Geman, Graffigne, & Dong 1990; Herault & Horaud 1993; Opara & Wörgötter 1998) and combinatorial optimisation applications (Peterson & Soderberg 1989; Blatt, Wiseman, & Domany 1997). Each feature is represented by a spin variable which attains one of a discrete set of spin states and a binding of two features corresponds to both sharing the same spin states. With regard to applications these models have the great advantage of being derived from energy or cost functions which characterize the stable output states as their minima. This energy-based approach establishes a link to pairwise clustering (Rose & Fox 1993; Hofmann & Buhmann 1997) and labeling problems in combinatorial optimisation (Kamgar-Parsi & Kamgar-Parsi 1990). Relaxation labeling (Rosenfeld, Hummel, & Zucker 1976), a standard technique in the field of pattern recognition, also falls into this category of energy-based labeling iteration schemes (Hummel & Zucker 1983; Pelillo 1994). Although offering a conceptual approach to binding, these models share certain drawbacks regarding their biological plausibility, since they either require iterative discrete cluster update procedures (Opara & Wörgötter 1998; Blatt, Wiseman, & Domany 1997) or complex normalizing nonlinearities (Peterson & Soderberg 1989; Rosenfeld, Hummel, & Zucker 1976; Hummel & Zucker 1983).

Since grouping is an intensively studied subject in computer vision, there exist a variety of other algorithms such as based on Markov Random Fields (Geman & Geman 1984), variational approaches (Mumford & Shah 1989), and curve evolution (Kimia, Tannenbaum, & Zucker 1995). See also (Wang & Terman 1997) for other related references. Recent work has stressed the importance of grouping for dealing with the occlusion problem (August, Siddiqi, & Zucker 1999; Elder & Zucker 1996). Another recent approach to segmentation has used normalized cuts (Shi & Malik 1997) to combine eigen-analysis with graph partitioning based on feature similarities.

In this contribution we analyse the competitive layer model (Ritter 1990; Wersing, Steil, & Ritter 1997) (CLM) which realises an energy-based approach to feature binding in a standard additive recurrent neural network with linear threshold neurons. Similar to spin models, a feature is assigned to one of a set of labels which are selected by a columnar local winner-take-all (WTA) circuit. These columns are coupled by lateral interactions which determine preferred bindings according to the mutual compatibility of the features. The neurons which represent the same label assignment are arranged in layers which are orthogonal to the columnar structure (see Fig. 1). As was shown in (Wersing & Ritter 1999), the attractors of the CLM provide feasible solutions to relaxation labeling problems. Ontrup & Ritter (1998) have applied the CLM to texture segmentation, based on features derived from Gabor filter banks, and shown that the method performs well on a wide range of image data.

The CLM energy function approach is similar to the quadratic energy function of Potts-Mean-Field (PMF) approaches to combinatorial optimization. An important difference is that the CLM is formulated as a piecewise linear system, and thus allows to use the tools of eigensubspace analysis for an inspection of the binding process. Our results extend previous results on single-column WTA networks (Sum & Tam 1996; Hahnloser 1998) to the layered multi-column case. By using dynamical WTA circuits as opposed to strict normalizations in PMF, the model gains more flexible and biologically plausible response properties. We show how this can be used advantageously in the grouping of contours and figure-ground segmentation.

In Section 2 we introduce the CLM architecture and discuss the main properties of its binding dynamics. Section 3 is devoted to a general theoretical analysis, which covers convergence, attractors and an eigensubspace analysis of the CLM with relation to the grouping properties. In Section 4 an efficient simulation procedure is stated, which we used for the application to contour grouping presented in Section 5. Section 6 discusses the results with respect to binding properties, noise tolerance and biological relevance.

2 The CLM feature binding model

2.1 The CLM architecture

The CLM consists of a set of L layers of feature-selective neurons (see Fig.1). We denote the activity of a neuron at position r in layer α by $x_{r\alpha}$ and denote as a *column* r the set of the neuron activities $x_{r\alpha}$, $\alpha = 1 \dots L$ that share a common position r in each layer. We associate each column with a particular "feature" which is described by a parameter vector $\mathbf{m}_r$. A typical feature example are local edge elements which are characterized by position and orientation, $\mathbf{m}_r = (x_r, y_r, \theta_r)$. More complex features were used by Ontrup & Ritter (1998) for texture segmentation, which employed a vector of local Gabor filter responses at different spatial frequencies and orientations.

A binding between two features, represented by columns r and r', is expressed by simultaneous activities $x_{r\hat{\alpha}} > 0$ and $x_{r'\hat{\alpha}} > 0$ that share a common layer $\hat{\alpha}$. Therefore, binding is achieved by having each (activated) column r assign its feature to one (or several, but see below) of the layers α, interpreting the activity $x_{r\alpha}$ as a measure for the certainty of that assignment. All the neurons in a column r are equally driven by an external input h_r which is to be interpreted as the significance of the detection of feature r by a preprocessing step. The afferent input h_r is fed to the activities $x_{r\alpha}$ with a connection weight $J_r > 0$.

Within each layer α the activities are coupled by the lateral interaction $\mathsf{f}^{\alpha}_{rr'}$ which corresponds to the degree of compatibility between features r and r' and which is a symmetric function of the feature parameters, $\mathsf{f}^{\alpha}_{rr'} = \mathsf{f}^{\alpha}(\mathbf{m}_r, \mathbf{m}_{r'}) = \mathsf{f}^{\alpha}(\mathbf{m}_{r'}, \mathbf{m}_r)$. The interactions $\mathsf{f}^{\alpha}_{rr'}$ determine which pattern configurations, if elicited as activity pattern within a single layer α, will be mutually supporting among their constituent parts ($\mathsf{f}^{\alpha}_{rr'} > 0$) or instead suffer mutual inhibition ($\mathsf{f}^{\alpha}_{rr'} < 0$). Two examples for lateral interactions motivated by Gestalt laws of perceptual grouping are shown in Figure 2. The lateral interaction pattern may be identical in all the layers or different to allow greater flexibility; a very useful case, which we dicuss for the purpose of perceptual grouping, is the incorporation of a "ground" layer to perform simultaneous grouping and figure-ground segmentation (see Sec. 5). The number of layers need not correspond to the number of groups, since for sufficiently many layers only those are active that carry a salient segment.

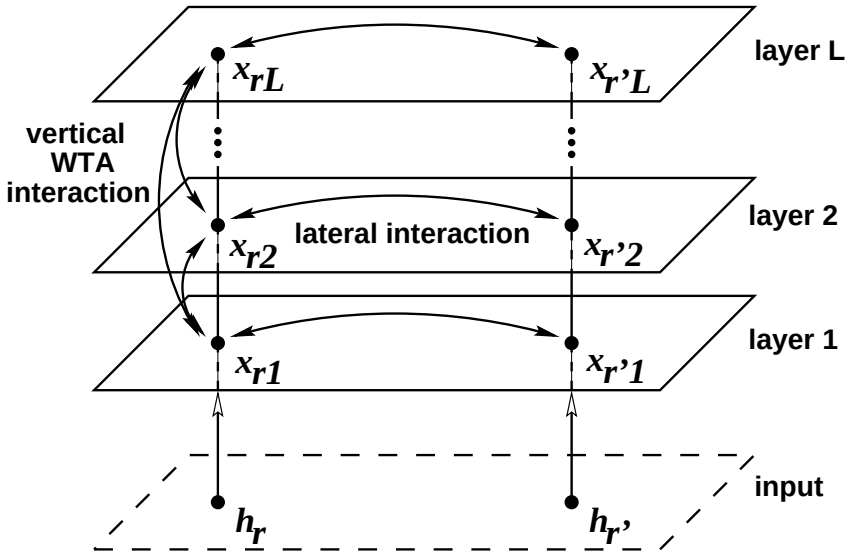


Figure 1: The CLM architecture. For each input feature, there is in each layer a responding neuron. A vertical WTA circuit implements a topographic competition between layers. Lateral interactions characterize compatibility between features and guide the binding process.

The purpose of the layered arrangement in the CLM is to enforce a dynamical assignment of the input features to the layers, using the contextual information stored in the lateral interactions. The assignment is realized by a columnar WTA circuit, which uses mutual symmetric inhibitory interactions with strength $I_r^{\alpha\beta} = I_r^{\beta\alpha} > 0$ between neural activities $x_{r\alpha}$ and $x_{r\beta}$ that share a common column r.

The combination of afferent inputs and lateral and vertical interactions can be combined into the standard additive activity dynamics

$$\begin{aligned}\dot{x}_{r\alpha} &= -x_{r\alpha} + \sigma\Big(J_r h_r - \sum_{\beta} I_r^{\alpha\beta} x_{r\beta} + \sum_{r'} \mathsf{f}_{rr'}^{\alpha} x_{r'\alpha} + x_{r\alpha}\Big) \\ &= -x_{r\alpha} + \sigma(E_{r\alpha} + x_{r\alpha}),\end{aligned} \tag{1}$$

where $\sigma(x) = \max(0, x)$ is a non-saturating linear threshold transfer function, and $E_{r\alpha} + x_{r\alpha}$ is the total input to neuron $x_{r\alpha}$. The additional self-excitatory term simplifies the following computations and can be compensated for by taking $\mathsf{f}'^{\alpha}_{rr'} = \mathsf{f}^{\alpha}_{rr'} - \delta_{rr'}$.

2.2 Energy Formulation and Binding Dynamics

The dynamics (1) has an energy function of the form

$$E = -\sum_{r\alpha} J_r h_r x_{r\alpha} + \frac{1}{2}\sum_{r}\sum_{\alpha\beta} I_r^{\alpha\beta} x_{r\alpha} x_{r\beta} - \frac{1}{2}\sum_{\alpha}\sum_{rr'} \mathsf{f}_{rr'}^{\alpha} x_{r\alpha} x_{r'\alpha}. \tag{2}$$

The energy satisfies $\partial E/\partial x_{r\alpha} = -E_{r\alpha}$, which makes E nonincreasing under the dynamics (1), $d/dt\, E = -\sum_{r\alpha} E_{r\alpha}(-x_{r\alpha} + \sigma(E_{r\alpha} + x_{r\alpha})) \leq 0$, since (1) confines the activities to be nonnegative (see App. 1). Thus the attractors of the dynamics (1) are the local minima of (2) under constraints $x_{r\alpha} \geq 0$. While the vertical interactions $I_r^{\alpha\beta}$ establish the columnar WTA process, the lateral interactions $\mathsf{f}_{rr'}^{\alpha}$ contribute in the quadratic energy as a sum over

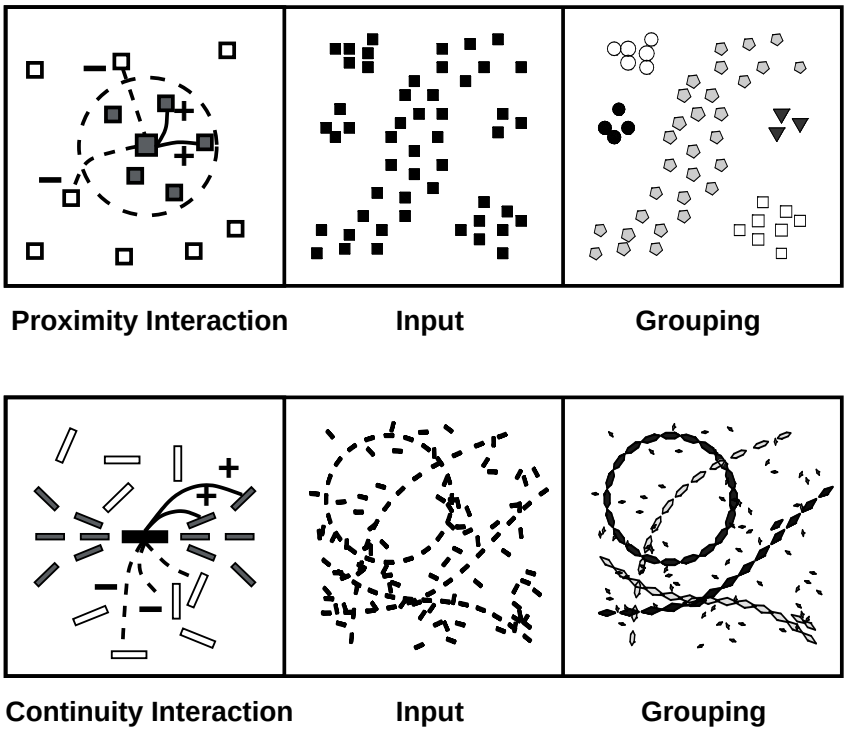


Figure 2: Examples of lateral interactions for perceptual grouping. An "OnCenterOffSurround" interaction which is excitatory for short distances and weakly inhibitory for larger distances is capable of clustering according to proximity of points. Different symbols denote activity in different layers. The Gestalt principle of continuity can be modelled by a local cocircular interaction of edge elements that lie along curves of constant curvature combined with a weak long-range inhibition to separate different segments.

all pairwise compatibilities within groups. If we interpret the negative value of the energy as the overall quality of the grouping, the aim of the dynamics (1) is to reach a globally minimal or almost minimal energy state.

The quadratic summation approach is similar to the corresponding Potts spin mean field free energy (Peterson & Soderberg 1989)

$$E_{PMF} = -\frac{1}{2}\sum_{\alpha}\sum_{rr'} \mathsf{f}^{\alpha}_{rr'} x_{r\alpha} x_{r'\alpha} + T\sum_{r\alpha} x_{r\alpha}\log x_{r\alpha}, \tag{3}$$

which is to be minimized subject to the weighting constraint $\sum_{\alpha} x_{r\alpha} = 1$. On the contrary, in a Potts spin model there is no explicit representation of an afferent input and the columnar activity must always sum to one due to its probabilistic interpretation. The additional convex entropy term that is weighted by the temperature parameter T biases the local minima of (3) towards soft columnar assignments, since it attains its minimum at $x_{r\alpha} = 1/L$. Minimal solutions to (3) can then be obtained by the recurrent dynamics

$$\dot{x}_{r\alpha} = -x_{r\alpha} + \frac{e^{F_{r\alpha}/T}}{\sum_{\beta} e^{F_{r\beta}/T}}, \qquad F_{r\alpha} = \sum_{r'} \mathsf{f}^{\alpha}_{rr'} x_{r'\alpha} \tag{4}$$

which implements a columnar WTA circuit that is gradually sharpened by decreasing (or "annealing") $T \to 0$.

The CLM dynamics (1) provides a similar soft competition scheme between all possible groupings in the different layers of the model. The hard weighting constraint and divisive nonlinearity (4), however, are replaced by a more flexible WTA circuit, that is coupled dynamically to the afferent input h_r and permits a contextual modulation of the input according to the salience of the grouping, but also requires an additional stability analysis. Perceptual context effects of this kind have been observed in a wide range of physiological (Gilbert 1992; Kapadia, Ito, & Westheimer 1995) and psychophysical (Kovács & Julesz 1993; Field, Hayes, & Hess 1992; Polat & Sagi 1994) studies. As our later eigensubspace analysis of the CLM dynamics reveals, the columnar WTA dynamics is driven by eigenmodes that depend on the lateral interaction pattern and whose eigenvalues characterize the rate of the group formation. The analysis motivates a mechanism for slowing down the lateral mode dynamics for increased grouping quality by adding a self-inhibitory coupling at each neuron of the form $\mathsf{f}'^{\,\alpha}_{rr'} = \mathsf{f}^{\alpha}_{rr} - T\delta_{rr'}$, where $T \geq 0$ is the strength of this self-inhibition. This leads to a new CLM energy

$$E' = E + T \sum_{r\alpha} x_{r\alpha}^2 \tag{5}$$

which adds an analog convex term that biases the local minima towards graded assignments and thus makes the WTA more soft. Similar to annealing in the Potts model (Peterson & Soderberg 1989), we can then use a gradual lowering of the inhibitory self-coupling T to sweep from graded assignments to the final unique assignment as the final grouping result.

A time course of the dynamics with and without self-inhibitory annealing performing contour grouping is shown in Figure 3. Annealing results in a much more regular group formation process, which produces groups proceeding hierarchically from more to less salient groups. The slowing down of the dynamics, however, results in a tradeoff for convergence time. Whether annealing is necessary depends on the complexity of the input pattern with respect to the lateral interactions. Two problems that have been emphasized by Wang & Terman (1997) are long-range coherence with local interactions and the proper separation of different segments. Self-inhibitory annealing increases long-range coherence, and also reduces the necessary strength of global inhibition for achieving separation, by a reinforcement of the corresponding dynamical modes, which we discuss in detail in the following section.

3 Theoretical Analysis

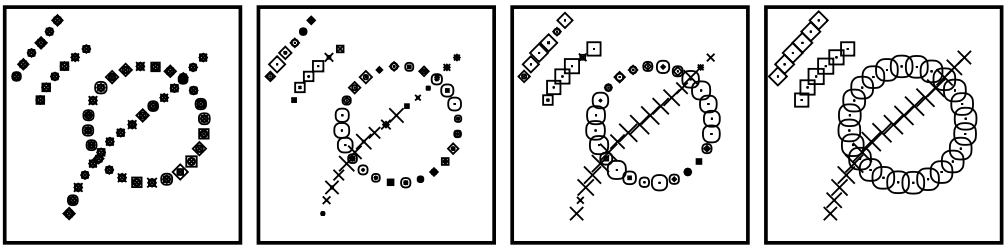
The dynamics (1) is a nonlinear dynamical system which consists of a series of continuously connected linear systems. An alternative dynamical implementation is of the form

$$\dot{x}_{r\alpha} = \omega_{r\alpha} E_{r\alpha} \tag{6}$$

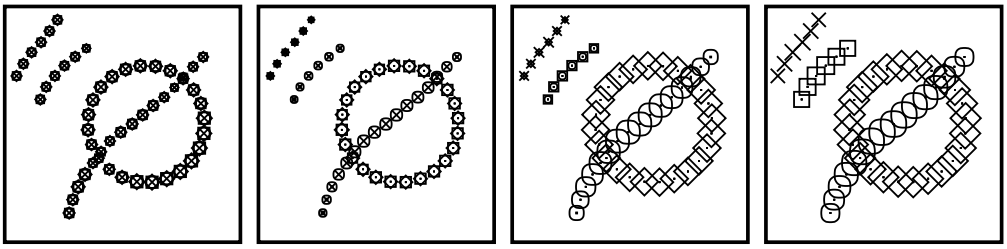
which avoids the occurrence of the $\sigma(.)$-nonlinearities and instead enforces the constraint $x_{r\alpha} \geq 0$ by a set of "switching variables"

$$\omega_{r\alpha} = \begin{cases} 0 & \text{for} \quad x_{r\alpha} = 0\,,\ E_{r\alpha} < 0 \\ 1 & \text{else} \end{cases} \tag{7}$$

This reformulation corresponds to constrained gradient descent in the energy (2) which naturally makes E nondecreasing, $d/dt\, E = -\sum_{r\alpha} \omega_{r\alpha} E_{r\alpha}^2 \leq 0$. Since both dynamics



a) Activity dynamics without annealing



b) Activity dynamics with self-inhibitory annealing

Figure 3: Activity dynamics and group formation in the CLM. The neural activity is displayed by different symbols for different layers with symbol size representing activity value. The lateral interaction is locally cocircular with a weak global inhibition as in Fig. 2. A typical time-course of the dynamics is shown in a). After initialization with random activity values first all activities within in a column are equally active in all layers. After that, dynamical modes, which break the symmetry between layers, cause formation of groups. The "soft competition" between the group assignments lasts until the columnar WTA circuit has caused an assignment to one of the layers. The modes are given by the eigenmodes of the lateral interaction pattern. The main problems are long-range coherence with local interactions and a proper separation of different groups. Fragmented groupings can occur especially for very short-ranged interactions due to fast expression of modes which lead to fragmented groups. b) shows how the mechanism of self-inhibitory annealing can be used to suppress these fragmenting modes. During the dynamics the strength of self-inhibition is lowered, which has the effect of making the columnar WTA circuit less strict and increases the "soft competition" between grouping assignments of a single feature. Fluctuations due to fragmented modes are smoothed out. The groups are then differentiated in a sequence, where the most salient structures appear first.

share the same energy function they converge to the same attractors which are local minima of (2) and may be considered as dynamically equivalent. Since the formulation (6) has the advantage of simplifying the linear analysis by shifting the nonlinearity to the boundary where $x_{r\alpha} = 0$, in the following we mainly consider the form (6). Nevertheless we prove our results in Appendix A in a form which also applies to the biologically more plausible and thus conventional form (1).

Our theoretical analysis is first based on a discussion of the model attractors and the conditions for convergence. We then discuss with an eigensubspace analysis which of the attractors are preferred by the recurrent dynamics and discuss with a synthetic example, how this relates to the properties of the binding process implemented by the CLM.

3.1 Convergence and Assignment

Networks composed of non-saturating linear threshold neurons as in (1) or (6) may be unbounded if the excitatory interactions are not balanced by sufficient inhibition. Hahnloser (1998) has discussed the stability for a single WTA circuit and given a criterion based on global inhibition. The following theorem, which we prove in the appendix, states that in fact local inhibition is sufficient to ensure boundedness for the CLM system of layerwise coupled WTA columns.

Theorem 1. *If $\kappa_{r\alpha} > 0$ with $\kappa_{r\alpha} = I_r^{\alpha\alpha} - \mathsf{f}_{rr}^{\alpha} - \sum_{r' \neq r} \max(0, \mathsf{f}_{rr'}^{\alpha})$, then the CLM dynamics is bounded. If $0 \leq x_{r\alpha}(0) \leq M$ for all r, α, where $M = \max_r (J_r h_r / \kappa_{r\alpha})$ then $0 \leq x_{r\alpha}(t) \leq M$ for all r, α and $t > 0$.*

The factors $\kappa_{r\alpha}$ control the stability margin and the maximal amplification of the inputs h_r. They are positive, if the self-interaction strength $I_r^{\alpha\alpha} - \mathsf{f}_{rr}^{\alpha}$ of a neuron is larger than the sum of the excitatory connections $\sum_{r' \neq r} \max(0, \mathsf{f}_{rr'}^{\alpha})$ converging onto it. If in the simplest case $J_r = I_r^{\alpha\beta} = J$ for all r, α, β, the afferent input and vertical interaction terms can be combined in the energy to the term $J \sum_r \left(h_r - \sum_\beta x_{r\beta}\right)^2$. We can interpret this as a penalty term with J as a constraint multiplier enforcing the condition, that the summed activity in a column is equal to the input h_r. Unlike in the usual penalty function approach to combinatorial optimisation problems, in the framework of sensory segmentation as we consider here a strict enforcement of this constraint is undesirable. If we choose J close to the stability margin of Theorem 1, the penalty term derived from the vertical interactions no longer dominates over the lateral interactions and the columnwise activity is strongly influenced by lateral effects; we return to this point after stating the next theorem.

Essential to the columnwise WTA approach is that we obtain a proper assignment of the features, that is, a state where each column r contains at most a single nonvanishing activity $x_{r\hat{\alpha}} > 0$. The conditions on the vertical interactions for this assignment property to hold are stated in the following theorem:

Theorem 2. *Let $F_{r\alpha} = \sum_{r'} \mathsf{f}_{rr'}^{\alpha} x_{r'\alpha}$ denote the lateral feedback of neuron r, α. If the lateral interaction is self-excitatory, $\mathsf{f}_{rr}^{\alpha} > 0$ for all r, α, and the vertical interactions satisfy $I_r^{\alpha\alpha} I_r^{\beta\beta} \leq (I_r^{\alpha\beta})^2$ for all α, β, then an attractor of the CLM has in each column r either*

i) at most one positive activity $x_{r\hat{\alpha}}$ with

$$x_{r\hat{\alpha}} = \frac{J_r}{I_r^{\hat{\alpha}\hat{\alpha}}} h_r + \frac{F_{r\hat{\alpha}}}{I_r^{\hat{\alpha}\hat{\alpha}}}, \quad x_{r\beta} = 0 \text{ for all } \beta \neq \hat{\alpha}, \tag{8}$$

where $\hat{\alpha} = \hat{\alpha}(r)$ is the index of the maximally supporting layer characterized by $F_{r\hat{\alpha}} > F_{r\beta}$ for all $\beta \neq \hat{\alpha}$, or

ii) all activities $x_{r\alpha}$, $\alpha = 1, \ldots, L$ in a column r vanish and $F_{r\alpha} \leq -J_r h_r$ for all $\alpha = 1, \ldots, L$.

This theorem states that as long as there are self-excitatory interactions within each layer, and the vertical cross-inhibition is sufficiently strong, the CLM converges to a unique assignment of features to the layer with maximal lateral feedback $F_{r\alpha}$. A central result is that, due to the layered topology of the network, this does not require arbitrarily large vertical couplings. Also note that the lack of an upper saturation is essential for the WTA behaviour, because it allows to exclude spurious ambiguous states. Nonzero activities are only stable due to a dynamical equilibrium and not due to saturation. Douglas et al. (1995) and Hahnloser (1998) have argued for the plausibility of this form of dynamical stability since cortical neurons rarely operate close to saturation. The lateral and vertical feedback modulates the input intensity by an amount that is dependent on the ratios of $I_r^{\alpha\alpha}, F_{r\alpha}$ and J_r. Consider again the simple example $J_r = I_r^{\alpha\beta} = J$, for all r, α, β where, according to Theorem 2, the output of the only active neuron in a column is then given by $x_{r\hat{\alpha}} = h_r + F_{r\hat{\alpha}}/J$. This shows that by lowering J the lateral effects can be increased and we obtain a context-dependent activity distribution, which still remains sensitive to the input intensities. Note that the additional self-inhibitory annealing by choosing $\mathsf{f}'^{\alpha}_{rr'} = \mathsf{f}^{\alpha}_{rr} - T\delta_{rr'}$ for T sufficiently large violates the condition of Theorem 2 and leads to stable states with multiple nonzero activity within a column, as the following eigensubspace analysis shows.

3.2 Eigensubspace Analysis

After we have discussed the possible attractors of the model, we now turn to the discussion which of them will be preferred by the dynamics. Since the CLM time development is defined by the piecewise linear differential equation (6), we can apply the method of eigensubspace analysis to obtain a solution in the linear domain, where the constraints are inactive. We consider the important special case of a lateral interaction $\mathsf{f}^{\alpha}_{rr'} = \mathsf{f}_{rr'}$ that is identical for all layers. For simplicity we take $J_r = I_r^{\alpha\beta} = J$ for all r, α, β as constant. As we will show in this section, the special form of two topological "orthogonal" interactions allows us to characterise the global eigenmodes completely in terms of the lateral eigenmodes.

Apart from constraints, the system is described by the linear dynamics

$$\dot{x}_{r\alpha} = Jh_r + \sum_{r'\beta} \mathsf{G}^{\alpha\beta}_{rr'} x_{r'\beta}, \tag{9}$$

where $\mathsf{G}^{\alpha\beta}_{rr'} = -J\delta_{rr'} + \delta_{\alpha\beta}\mathsf{f}_{rr'}$. Introducing the $N \times L$ vectors

$$\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_L) \quad \text{with} \quad \mathbf{x}_\alpha = (x_{1\alpha}, \ldots, x_{N\alpha}) \quad \text{and} \tag{10}$$

$$\mathbf{h} = (\mathbf{h}_0, \ldots, \mathbf{h}_0) \quad \text{with} \quad \mathbf{h}_0 = (h_1, \ldots, h_N) \tag{11}$$

we can write this as

$$\dot{\mathbf{x}} = J\mathbf{h} + \mathsf{G}\mathbf{x} \tag{12}$$

where G is a linear transformation in the space $\mathbb{R}^{N\times N} \otimes \mathbb{R}^{L\times L}$ and can be decomposed as

$$\mathsf{G} = \mathsf{f} \otimes \mathsf{Id}^{L\times L} - J\, \mathsf{Id}^{N\times N} \otimes \mathbb{I}^{L\times L} \tag{13}$$

with Id as the identity and $\mathbb{I}^{L\times L}$ as an $L \times L$-matrix of 1's. Equation (13) shows that an orthonormal eigenvector basis $\{\mathbf{v}^{i\gamma}, \Lambda_{i\gamma}\}$ for G can be obtained from orthonormal eigenvector bases [1] $\{\mathbf{b}^i, \lambda_i\}_{i=1\ldots N}$ and $\{\mathbf{q}^\gamma, \mu_\gamma\}_{\gamma=1\ldots L}$ for f and $\mathbb{I}^{L\times L}$ respectively:

$$\mathbf{v}^{i\gamma} = \mathbf{b}^i \otimes \mathbf{q}^\gamma, \qquad i = 1\ldots N,\ \gamma = 1\ldots L, \tag{14}$$

$$\Lambda_{i\gamma} = \lambda_i - J\mu_\gamma. \tag{15}$$

$\mathbf{q}^\gamma, \mu_\gamma$ can be calculated analytically. Here, we only use that the first eigenvector obviously is given by

$$\mathbf{q}^1 = 1/\sqrt{L}(1,\ldots,1), \qquad \mu_1 = L, \tag{16}$$

while the remaining orthogonal eigenvectors have zero eigenvalues $\mu_2 = \mu_3 = \cdots = \mu_L = 0$. This has a direct geometrical interpretation. We can divide the eigenmodes of the linear system into two classes:

- DC-Subspace: This is the subspace spanned by the eigenmodes

  $$\mathbf{v}^{i1} = 1/\sqrt{L}(\mathbf{b}^i,\ldots,\mathbf{b}^i), \quad \Lambda_{i1} = \lambda_i - JL \tag{17}$$

  which contains the coherent eigenmodes which have equal components in all layers. For sufficiently large J all the corresponding eigenvalues Λ_{i1} are negative.

- AC-Subspace: This is the orthogonal subspace spanned by the remaining eigenvectors

  $$\mathbf{v}^{i\gamma\neq 1} = (q_1^\gamma \mathbf{b}^i,\ldots,q_L^\gamma \mathbf{b}^i), \quad \Lambda_{i\gamma} = \lambda_i \tag{18}$$

  which contain the eigenmodes causing differences in the activity patterns between layers and which have a zero mean summed over columns. Their eigenvalues $\Lambda_{i\gamma\neq 1}$ are given by the eigenvalues λ_i of the lateral interaction $\mathsf{f}_{rr'}$ alone and are $L-1$ degenerate due to the symmetry between layers.

The linear dynamics can be completely characterized by the eigensubspaces and the location of the fixed point which is determined by the inhomogeneous input $\mathbf{h}$. Relative to the fixed point, the components $c_{i\gamma}(t) = (\mathbf{x}(t) - \mathbf{x}^F)\cdot\mathbf{v}^{i\gamma}$ develop according to $c_{i\gamma}(t) = c_{i\gamma}(0)e^{\Lambda_{i\gamma}t}$. This corresponds to exponential growth for positive eigenvalues, while negative eigenvalues lead to attracting affine constraint surfaces which pass through the fixed point.

By expanding the fixed point equation $\dot{\mathbf{x}} = 0$ in the basis $\mathbf{v}^{i\gamma}$ we obtain the following expansion of the fixed point $\mathbf{x}^F$ in powers of λ_i/JL (note that $\mathbf{h}$ and thus also $\mathbf{x}^F$ has no component in the AC space):

$$\mathbf{x}_\alpha^F = \frac{1}{L}\mathbf{h}_0 + \frac{1}{L}\sum_i \frac{\lambda_i}{JL}(\mathbf{b}^i\cdot\mathbf{h}_0)\mathbf{b}^i + \mathcal{O}\Big(\frac{\lambda_i}{JL}\Big)^2. \tag{19}$$

Therefore, if $JL \gg \lambda_i$ we can approximate the fixed point by $\mathbf{x}_{r\alpha}^F = h_r/L$. We can now draw a sketch (see Fig.4) of the initial CLM dynamics referring to our eigensubspace analysis.

[1] We assume that the λ_i are sorted in descending order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$

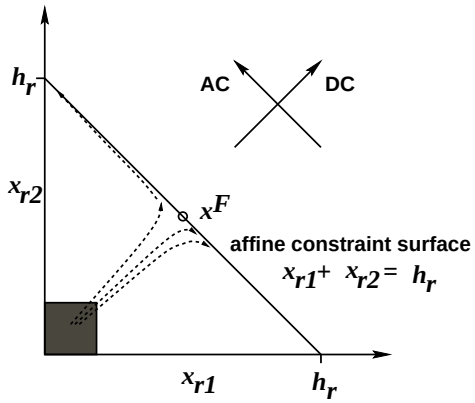


Figure 4: Sketch of the linear dynamics for two layers. Shown are the activity trajectories for the two activities x_{r1}, x_{r2} of a single column r. Starting from small initial values (grey square) the activities quickly approach the fixed point $\mathbf{x}^F$ and the constraint surface $\sum_{\alpha=1,2} x_{r\alpha} = h_r$ in the DC subspace. Then the dynamics in the orthogonal AC subspace drives the WTA process until only one layer is active.

Suppose $h_r > 0$ for all r and we initialise the system with small positive random values $0 < x_{r\alpha}(0) \ll h_r$. For sufficiently large $JL \gg \lambda_i$ all eigenmodes in the DC-subspace have large negative eigenvalues. Therefore the initial state will be rapidly projected onto the affine subspace $\sum_\alpha x_{r\alpha} = \sum_\alpha \mathbf{x}^F_{r\alpha}$ orthogonal to the DC-subspace which passes through the fixed point. The rapid projection dynamics in the DC space moves the activities away from the zero boundary and thus keeps the initial dynamics linear. Now, the modes in the AC-subspace with positive eigenvalues λ_i develop differences between layers on a slower timescale, since their eigenvalues have smaller absolute value. The principal AC-mode with the largest positive eigenvalue dominates the AC-subspace dynamics in this initial phase and determines the timescale of the WTA dynamics which drives some activities to zero in a column while others are increased. After some of the activities reach zero level, they no longer contribute to the linear dynamics and the nonlinearity is taking effect by introducing a new segment of the piecewise linear dynamics with different resulting eigenmodes.[2] Since then the symmetry between layers is broken, the coupled WTA dynamics gets more complex in the general case, by subsequently driving the unassigned columns towards an assigned state as can be seen by comparison with Figure 3. For regular and highly symmetric patterns, however, which we discuss in the following section, the final grouping result can be already obtained from the dominant mode in the initial linear phase of the dynamics.

The suggested mechanism of self-inhibitory annealing by choosing $\mathsf{f}'^{\,\alpha}_{rr'} = \mathsf{f}^\alpha_{rr} - T\delta_{rr'}$ shifts all eigenvalues by $\lambda'_i = \lambda_i - T$. If T is greater than the critical value $T_c = \lambda_1$, then all eigenvalues are negative and the unassigned fixed point $\mathbf{x}^F$ is asymptotically stable, which is analogous to Potts mean field models (Peterson & Soderberg 1989). If then T is lowered from the critical value, at first only the $\mathbf{b}^1$-AC modes acquire a positive eigenvalue and grow exponentially away from the fixed point. The following section discusses how this can be employed to increase the grouping quality.

[2]The resulting interaction matrix is of the form $\omega_{r\alpha}\omega_{r'\beta}\mathsf{G}^{\alpha\beta}_{rr'}$.

3.3 Coherence, Separation, and Self-Inhibitory Annealing

Let us consider two simple stimulus examples for the contour grouping lateral interaction to discuss the binding problems of long-range coherence and separation referring to the previous eigensubspace analysis. The first stimulus consists of a circle, composed of N regularly placed edge elements (see Fig. 5). We assume that we can neglect the effect of all other columns r' which are at subthreshold activity (i.e. $h_{r'} < 0$), and therefore consider the lateral eigenvectors only from the resulting $N \times N$ lateral interaction matrix $\mathsf{f}_{rr'}$ of the active features. The contour grouping lateral interaction has two contributions, the local cocircular interaction $f^{\text{cocirc}}_{rr'}$ (see App.2) and a global inhibition with strength k, giving together $\mathsf{f}^{\alpha}_{rr'} = f^{\text{cocirc}}_{rr'} - k$. Since the overall interaction pattern is rotationally invariant on the circle, the eigenvectors of the interaction are the sine and cosine waves on the circle with corresponding spatial frequencies.

We first assume $k = 0$. Each edge shares positive interactions in a local neighbourhood range and the zeroth order (constant) cosine with $b^1_r = 1$ is the eigenvector with largest eigenvalue $\lambda_1 = \sum_{r'} f^{\text{cocirc}}_{rr'}$, equal to the sum over the local cocircularity values. The corresponding global AC modes coherently increase and decrease the activity on the circle in the different layers. The next two modes with degenerate eigenvalues $\lambda_{2,3} < \lambda_1$ are the first order sine and cosine waves, which split the circle into two halves. The gap between λ_1 and $\lambda_{2,3}$ determines whether the coherent mode dominates over the fragmenting modes. It is a general property of local interactions that this gap tends to zero, if the range of the interactions tends to zero.[3] Since the difference between the modes grows exponentially, also the time the system stays in the initial linear mode determines the degree of coherence. By shifting the spectrum of the lateral interaction with the self-inhibitory addition $\mathsf{f}'_{rr'} = \mathsf{f}_{rr'} - T\delta_{rr'}$, this time is increased. This can be used to dynamically suppress the spurious $\mathbf{b}^{2,3}$ modes, however also reduces the overall convergence rate.

Now suppose we add a second equal and well-separated circle, such that there are no local cocircular interactions between the circles (see Fig. 5). Now the task is not only to achieve coherence on the circles, but also proper separation. Since the lateral interaction pattern is exchange-symmetric between the circles, the eigenvectors of the composite system are symmetric and antisymmetric combinations of the single circle eigenvectors with the same eigenvalues. The modes $\mathbf{b}^1$ and $\mathbf{b}^2$, responsible for separating and grouping together the circles, respectively, are degenerate with $\lambda_1 = \lambda_2$ without any global inhibition. If we now add the global lateral inhibition by choosing $k > 0$, only the $\mathbf{b}^2$ mode undergoes a change in eigenvalue, since all other modes have a zero component within this global constant inhibition. The effect of k is then given by shifting the eigenvalue λ_2 of the spurious mode to the value $\lambda_2 - 2kN$. Therefore, if k is sufficiently large, the mode which separates the two circles coherently is dominant. There is, however, an obvious limitation to increasing k, since a strong global inhibition destabilizes the coherence within a circle. Similarly, for the single circle the inhibition lowers the eigenvalue of the single coherent mode by $\lambda'_1 = \lambda_1 - kN$. Therefore, to achieve both long-range coherence and separation only small values of k are possible, for our applications we choose $k = z/N$, where $0 < z \leq 1$.

The picture gets considerably more complex if there are more complex interacting

[3] Suppose the spatial exponential decay of the cocircular interaction is replaced by a hard cut-off function, such that the interaction of two perfectly cocircular features on the circle at angular positions $\phi_r, \phi_{r'} \in [0, 2\pi]$ is $\mathsf{f}_{rr'} = 1$ if $|\phi_r - \phi_{r'}| < D$ and $\mathsf{f}_{rr'} = 0$ otherwise. The difference between the coherent mode $b^1_r = 1$ and the first fragmenting modes $b^2_r = \sin(\phi_r), b^3_r = \cos(\phi_r)$ is in the limit of sufficiently small feature spacing given by $2(D - \sin(D)) \to 0$ for $D \to 0$.

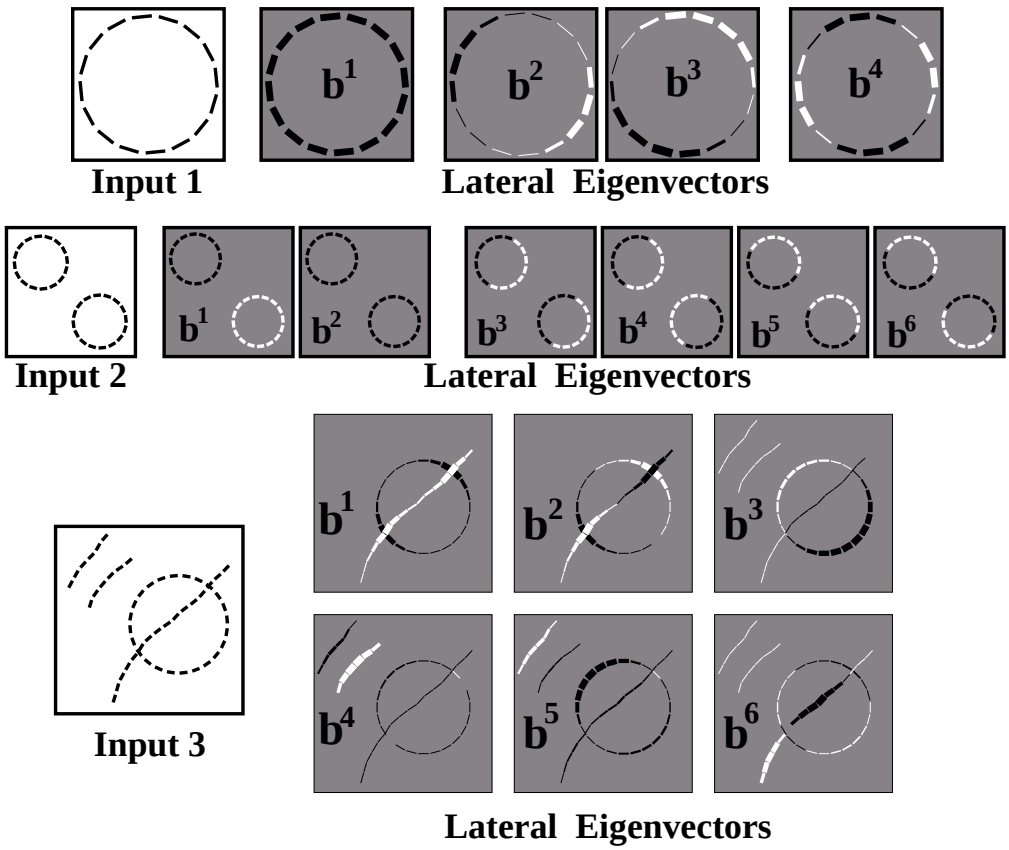


Figure 5: Lateral eigenvectors of the contour grouping interaction. The top row shows the lateral eigenvectors of the lateral interactions on the active features of this circle input pattern. Vector components are shown as black (positive) and white (negative) edges with thickness proportional to magnitude. To achieve a coherent binding on the circle, the eigenvalue λ_1 of the coherent mode $\mathbf{b}^1$ must be sufficiently large compared to the eigenvalues $\lambda_{2,3}$ of the circle-splitting modes $\mathbf{b}^{2,3}$. For two well-separated circles (middle row) the eigenvectors are symmetric and antisymmetric combinations of the single circle eigenvectors. Without global inhibition, the modes $\mathbf{b}^1$ and $\mathbf{b}^2$ are degenerate. A weak global inhibition suffices to lower the eigenvalue λ_2 and thus suppress the unwanted $\mathbf{b}^2$ mode. For more complex inputs (lower row) the succession of eigenvectors constitutes a multi-scale hierarchy. The principal eigenmode $\mathbf{b}^1$ separates the circle from the crossing line and the short lines. $\mathbf{b}^4$ and $\mathbf{b}^5$ drive the segregation of the short lines. The simple self-inhibitory annealing scheme proposed in section 3.3 can be interpreted as hierarchically suppressing subsequent eigenmodes resulting in a hierarchical group formation process.

groups in the feature input as the lower example in Fig.5 shows. The sign-distribution of the lateral eigenvectors constitutes a multi-scale hierarchy of the structures in the input. By gradually lowering the self-inhibition T, which can be understood as a type of annealing process with regard to the energy function of the dynamics, these modes are expressed in a hierarchical sequence, as can be seen by looking back at Fig.3.

4 Simulation

The CLM dynamics can be simulated in principle in parallel by any differential equation integrator like the Euler or Runge-Kutta method. Also, due to the piecewise linearity, an analog VLSI implementation may be even simpler than classic circuits incorporating sigmoid transfer functions. If simulated on a serial processor, there is an alternative approach which replaces the explicit trajectory integration by a rapid search for fixed point attractors. This can be done by iteratively solving the fixed point equations which is largely facilitated by the piecewise linearity of the model (Ontrup & Ritter 1998). This iterative solution procedure, also known as a Gauss-Seidel approach has also been extensively used for Markov Random Field approaches to image segmentation (Besag, Green, Higdon, & Mengersen 1995). The algorithm in conjunction with an exponential annealing schedule for the temperature parameter T can be implemented in the following way:

1. Initialize all $x_{r\alpha}$ with small random values around

   $$x_{r\alpha}(t=0) \in [h_r/L - \epsilon, h_r/L + \epsilon].$$

   Initialize T with $T = T_c$.

2. Do $N \cdot L$ times: Choose (r, α) randomly and update $x_{r\alpha} = \max(\,0,\,\xi\,)$, where

   $$\xi = \frac{J_r h_r - \sum_{\beta \neq \alpha} I_r^{\alpha\beta} x_{r\beta} + \sum_{r' \neq r} \mathsf{f}^{\alpha}_{rr'} x_{r'\alpha}}{I_r^{\alpha\alpha} - \mathsf{f}^{\alpha}_{rr} + T} \quad .$$

3. Decrease T by $T := \eta T$, with $0 < \eta < 1$. Go to step 2 until convergence.

The single activity update in step 2 corresponds to solving the fixed point equation for this activity with all other activities held constant. For the figure-ground setup as described in section 2 the critical temperature is given by $T_c = \lambda_{max}\{\mathsf{f}_{rr'}\}$. This asynchronous dynamics converges (Ontrup & Ritter 1998) due to a convergence result on asynchronous iteration in neural networks by Feng (1997). We observed a speed gain of roughly a factor of ten compared to Euler integration at comparable solution quality. For simulation without self-inhibitory annealing simply set $T = 0$.

5 Application to Contour Grouping

5.1 Related Work

It is long known that in the early stages of visual information processing in visual cortex area V1 many neurons can be found (Hubel & Wiesel 1962) that respond to local oriented edge

elements within their classical receptive field. One of the major issues in vision research are the mechanisms which are used by the visual system to integrate these local elements into global salient contours to facilitate robust boundary detection and object recognition. This process has been considered (Sajda & Finkel 1994; Zucker, Dobbins, & Iverson 1989) as being composed of two components: the process of enhancement, where local edge information is combined cooperatively for smooth and salient contours, and the process of segmentation or grouping, where separate contour elements are assigned to different groups.

Recent neural models of contour integration have mainly focused on the enhancement stage of visual processing. Many models of orientation tuning are composed of competitive orientational "OnCenterOffSurround" interactions within hypercolumns (Ben-Yishai, Lev Bar-Or, & Sompolinsky 1995; Mundel, Dimitrov, & Cowan 1997), that are modulated by experimentally found horizontal or lateral long range interactions (Gilbert 1992; Weliky, Kandler, Fitzpatrick, & Katz 1995) to produce contextual effects. Enhancement through feature linking is addressed by the model of Yen & Finkel (1996, Yen & Finkel (1998), which, however, requires rule-based interactions and global activity normalizations. Li (1998) has stated a more biologically plausible dynamical model for contour integration in area V1 and discusses properties of the oscillatory correlations in the model which may be useful to facilitate synchronization-based segmentation. As the results of Li (1998) demonstrate, the complex excitatory and inhibitory interactions lead to distorted and less predictable correlations for more complex visual scenes, leaving the question of a robust grouping process mostly unanswered.

In the following we present the application of the CLM as a new approach to the grouping stage of contour integration. To keep the model simple and to allow for the application to large feature sets of real images, we do not consider the generation of local orientation within a hypercolumn and leave this to a preprocessing step. The essential WTA then does not operate between local orientation alternatives within a hypercolumn, but between different grouping alternatives of the edge feature with fixed local orientation.

5.2 The CLM Contour Grouping Model

We suppose that in an initial integration step the local edge information has been subsampled and preprocessed. As a result of the preprocessing, we assume that within a localized area only a single edge detector is active at an optimally tuned orientation. We therefore consider a set of idealized edge feature detectors, indexed by $r \in \{1, \ldots, N\}$ and characterized by position, local orientation and an associated edge intensity value h_r.

The CLM Contour Grouping Model is then composed of a set of L layers $\alpha \in 1, \ldots, L$ where $L - 1$ "figure" layers are provided to respond to salient contour groups and a single "ground" layer is provided to capture the background features. The pairwise lateral interaction between edge features in the figure layers is visualized in Figure 6. The excitatory component links edges that are co-circular, i.e. lie tangentially to a common circle, with a tolerance that admits small deviations. This orientational field is superimposed with a Gaussian distance-dependent component. This excitatory interaction field is similar to recent models (Parent & Zucker 1989; Yen & Finkel 1998; Li 1998). A similar interaction pattern between orientation sensitive cells has been observed experimentally (Gilbert, 1992; Weliky et al., 1995). The local inhibitory component is of comparable magnitude as the excitatory field and insensitive to orientation as suggested by experimental findings (DeAngelis et al., 1994). To support the separation of remote contour parts an additional weak long-range inhibition is necessary, which has no direct biological counterpart in the

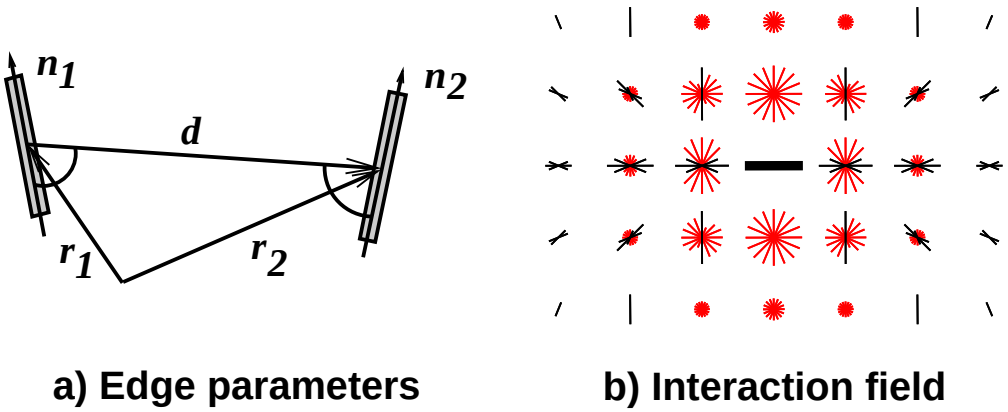


Figure 6: Lateral interaction for contour grouping. a) The pairwise interaction depends on difference vector $\mathbf{d}$ and two unit vectors $\mathbf{n}_1, \mathbf{n}_2$ encoding orientation for computational convenience. b) shows the resulting interaction pattern for a single horizontal edge at the center. Sampled at surrounding positions and orientations, black edges share excitatory, grey edges share inhibitory interaction with the central edge. Length codes for interaction strength.

experimentally observed lateral connection structure.[4] The complete lateral interaction, given as $f^{\text{cocirc}}_{rr'}$, and the preprocessing stage which we employed for the feature generation are defined in Appendix B.

The lateral interactions $\mathsf{f}^{\alpha}_{rr'}$ are then given as $\mathsf{f}^{1}_{rr'} = m\delta_{rr'}$ for the ground layer and as $\mathsf{f}^{\alpha>1}_{rr'} = f^{\text{cocirc}}_{rr'} - k$ in the figure layers. The parameter $m > 0$ defines a self-coupling against which lateral interactions in the figure layers must compete to "pop out" a feature from the ground layer. The weak global inhibition strength k was chosen as $k = 0.3/N$. The vertical interactions and input strength are chosen equally as $J_r = I_r^{\alpha\beta} = J$ with $J > J_c = \max_r \sum_{r'} \sigma(\mathsf{f}_{rr'})$, where J_c is the critical value implied by the stability margins of Theorem 1. We used a value of $J = 1.1 J_c$.

The application of the CLM contour grouping model to a real image is shown in Figure 7. The preprocessing (see App. B) results in a set of features where object and background textures produce a noisy background. To obtain a robust grouping for the complex scene of interacting edge elements, annealing in T was necessary, where we used an exponential schedule as described in Section 4 with $\eta = 0.99$. The grouping result is shown in Fig. 7c. By an appropriate choice of the ground layer strength $m = 3.5$, which must approximately match the sum over the cocircular interactions of an edge that is part of a proper contour, the salient contours are detected as single groups. Since noisy edges lack the lateral support of aligned edges, they are then captured by the ground layer and for sufficiently many available layers only those containing salient segments will be active. The lateral interactions cause an amplification of salient low-intensity contours without enhancing noisy fragments. To demonstrate that this amplitude-dependent modulation is crucial in the model, we compared the CLM grouping result to a Potts spin mean field model with equal lateral interaction structure and annealing. Spin models generally lack the ability to represent additional amplitude information. The result in Fig. 7d illustrates that this makes them more sensitive to noise in the image, because noisy aligned elements may form spurious groups. Since we

[4] Wang & Terman (1997) have speculated about the thalamus providing the global inhibition.

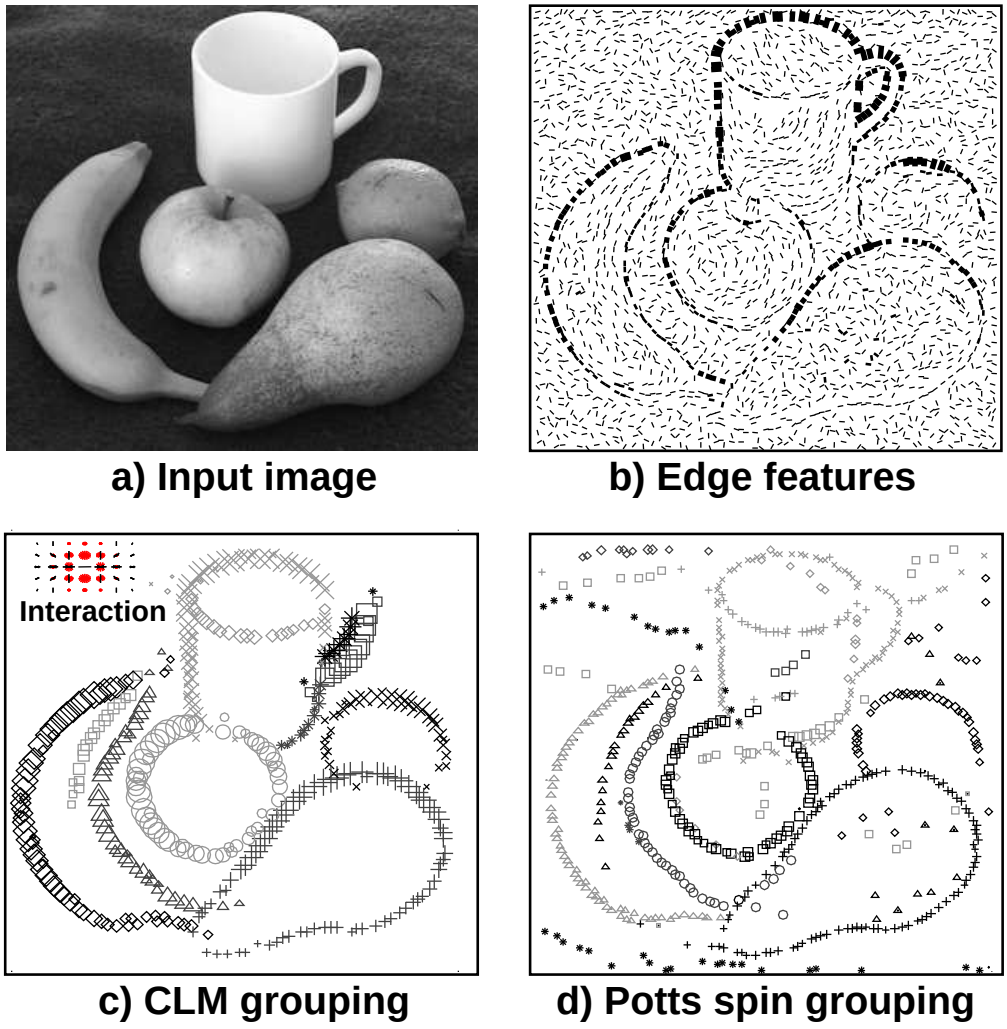


Figure 7: Grouping of a natural image. An input image a) is preprocessed to generate a set of ≈ 2000 edge features displayed in b). The edge input intensity h_r is visualized as the thickness of the displayed elements. The noise from the background and fruit textures has low amplitude, however, cannot be suppressed by thresholding without losing low contrast contours of the objects. c) shows the grouping result for a CLM architecture with 20 figure layers and one ground layer. The symbols (7 symbols $\times$ black,light, and dark grey) represent activity in different layers with symbol size proportional to magnitude. The ground layer is omitted. The scale of the lateral cocircularity interaction is displayed in the upper left corner. The CLM grouping gives 11 segments, which achieve an identification of the most important curve elements in the presence of background noise, where some of the layers remain inactive. The low-intensity edges of the pear and apple are amplified due to the supporting lateral interaction on the salient contours. d) shows the grouping by a Potts spin system with the same interactions. Due to the lack of amplitude information the model has a strong tendency to "hallucinate" curves which cannot be compensated by the ground layer.

were not able to produce significantly better results also by varying parameters m and k we conclude that the amplitude-dependent dynamics in the CLM leads in this setting to a better signal-noise separation for complex input patterns.

Figure 8 shows the performance of the CLM with the same parameter settings and identical annealing on more complex image data, where for an aerial image the number of layers is not sufficient to allow a single layer for each salient segment. In that case the dynamics tries to find a compromise by combining segments into the same layer such that their mutual inhibition is minimal and thus the overall energy level is minimal. This results in a tendency to prefer a combination of parallel segments, which are separated by a distance larger than the local inhibition field of the cocircular interaction. In principle, it would be possible to suppress multiple segments within a layer by raising the global inhibition. This, however, in turn poses an upper limit on the maximum size of a stable segment, since then the global inhibition may exceed the local cocircular support. The time course of the dynamics during self-inhibitory annealing illustrates the resulting hierarchical group formation process which initially expresses the most salient groups in the available layers.

6 Discussion

6.1 Binding Properties

The CLM binding model shares some components with other recent binding models. For the contour grouping example that we have considered, the lateral interactions are based on a local compatibility combined with a weak global inhibition, which is the same for LEGION (Wang & Terman 1997) and the ECU spin update model (Opara & Wörgötter 1998). The LEGION model also uses a mechanism to perform figure-ground segmentation by suppressing features with low lateral support. The main difference lies in the dynamical implementation, which for the CLM is given by a consistent model of neural activity dynamics in a layered system of coupled WTA columns. For the local contour grouping model, long-range coherence and separation are enforced by a self-inhibitory annealing mechanism, that causes controlled expression of dynamical modes, but also results in a tradeoff for convergence time. Another CLM application to texture segmentation (Ontrup & Ritter 1998), has shown that for sufficiently dense and long-ranged feature interactions, the CLM performs well also without annealing. The models of Wang & Terman (1997) and Opara & Wörgötter (1998) employ mechanisms similar to region growing and have been successfully applied to the task of greyscale segmentation of image data, however we believe that the complex interactions of overlapping and intersecting curve elements might pose problems for their application to contour grouping.

Our eigensubspace analysis reveals an interesting parallel between the grouping dynamics in the CLM and the oscillatory synchronizations in the contour integration model of Li (1998). If we restrict ourselves to lateral interactions which are only excitatory, then the dominant oscillation mode in that model has the same eigenvector components as the dominant grouping mode in the CLM. Therefore, a similar mechanism of suppressing spurious eigenmodes could be used to improve the synchronization properties in that model. We consider a combination of synchronization-based and spatial mechanisms as a promising approach to more realistic models of cortical feature binding, which have enough computational power to be used for sensory segmentation tasks.

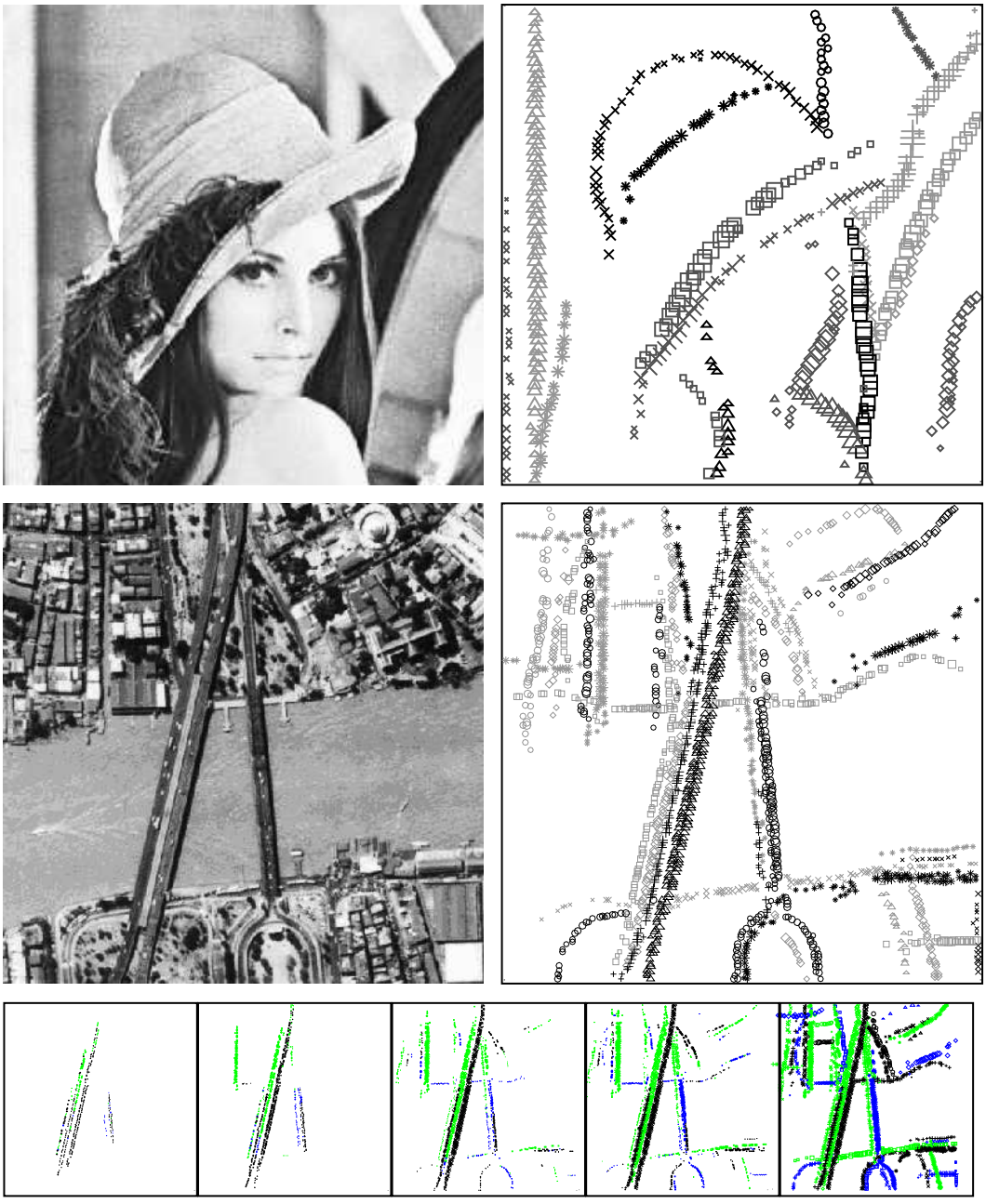


Figure 8: Grouping Results. The figure shows the CLM grouping results for two other images with identical parameter settings and 20+ground layers. For the "leena" image ($\approx$ 2000 features) the CLM gives 18 segments which cover the most salient contours. The ground layer strength introduces a minimum size that a segment must have to collect enough lateral support, thus smaller contour segments are not resolved and would require a multi-scale extension of the model. Low intensity edges are enhanced at brim and the top of the hat. The lower aerial image (≈ 3500 features) is taken from (Yen & Finkel 1996). Due to the presence of more salient contours than layers, some layers actually carry more than a single segment, like e.g. the "black circle" layer, which has three vertical groups in the upper left corner and the center of the image. This illustrates a tendency to combine parallels into the same layer which is most efficient with regard to the energy level of the global grouping. The lower row shows the sequence of emerging groups when lowering the self-inhibition from left to right.

6.2 Selective Amplification and Noise Tolerance

A main difference of the CLM compared to spin models of segmentation is the explicit usage of amplitude information in the input. The usage of dynamical WTA circuits results in a context-dependent amplitude modulation which enhances salient groups due to their supporting lateral interactions. Our comparison to a Potts spin model shows that this results in a better noise tolerance for the contour grouping example. This process has been discussed by Li (1998) and Li & Dayan (1999) as selective amplification. We have considered a setup where only the contributions of a set of coarsely sampled features are active, while others are omitted from the model. This decreases the ability to hallucinate contours, since we do not have a full range of feature detectors at all orientations as in the model by Li (1998). This also excludes a direct "filling in" of contours which are not present in the input. To ensure a proper operation of the model in an extended case of features at all orientations, it would be necessary, similar as in (Li 1998), to increase the threshold of the non-saturating linear transfer function from zero to a finite positive value to suppress the activation of neurons with zero afferent input.

6.3 Biological Relevance

The key principle of the CLM architecture is the encoding of feature bindings by the assignment to separate populations of laterally interconnected and locally competitive neurons. A biologically realistic interpretation of the CLM architecture may be expressed either in terms of the prominent layered structure of the real visual cortex, or it may be implemented in the rich local connectivity structure of a single neuronal layer itself. Here, we want to restrict our discussion on the macaque monkey primary visual cortex, area V1, which is one of the most intensively studied regions of the brain.

Although it is known that neurons in the granular and supergranular layers of the primary visual cortex are connected via horizontal connections of considerable lateral spread (Rockland & Lund 1983) there is currently no experimental data available which supports the idea of specific inhibitory (competitive) interactions between disjunct modules of laterally interconnected neurons. However, the CLM model illustrates, that this principle could be easily implemented in a simple neural circuit and therefore we propose to look for related anatomical structures as an interesting experimental question. The presence of sufficiently many layers to carry each salient segment in an image is hard to justify in the visual system, albeit it may be useful from a technical applications point of view. Nevertheless, since our suggested self-inhibitory control mechanism together with the figure-ground separation allows for a concentration on the most salient groups, a limited number of layers may not form a severe restriction.

The principle of feature binding within competitive and spatially segregated layers may help to process visual information via different routes through the cortical neuropile. For example, there is evidence of local topographic inhibition between those layers in area V1 which receive the majority of afferent inputs from the LGN (Hubel & Wiesel 1972); the principal input layers $4C\alpha$ and $4C\beta$ are coupled by reciprocal inhibitory interneurons (Lund 1987) (varieties α-3, β-3). Since the α and β sublayers are characterized by different response properties (Blasdel & Fitzpatrick 1984; Hawken & Parker 1984) as well as different spread of lateral connections (Yoshioka, Levitt, & Lund 1994), a putative interpretation of the inhibitory circuitry could be that feature binding in each sublayer is performed by such a competitive interaction. Moreover, the neurons in layer $4C\alpha$ and $4C\beta$ project to different

target layers within area V1, which in turn contain key sets of efferent neurons to other cortical areas (Yoshioka, Levitt, & Lund 1994; Yabuta & Callaway 1998). Therefore, a binding of features to different depth of layer 4C may form the basis of the subsequent processing stages.

The layered structure of the CLM involves a certain degree of neural redundancy and could be considered as a "waste" of neural hardware. The fixed hardwired layers might seem less flexible than the synchronization-based approach, which carries additional information into the temporal domain. Models have shown, however, that complex interactions lead to strong limitations with regard to stability and separation of the groupings. We suggest the principle of topological segregation, as used in the CLM, as an additional binding principle which could improve the robustness and stability of the feature binding processes.

Acknowledgements

The authors would like to thank Ute Bauer and Wolf Jürgen Beyn for helpful discussions. We also thank the anonymous referees for useful comments which improved the clarity of the manuscript. H. Wersing is supported by DFG grant GK-231. J. J. Steil is supported by DFG grant Ri-621-2-1.

Appendix A: Proof of Theorems

We consider the CLM dynamical system as

$$\dot{x}_{r\alpha} = \omega_{r\alpha} E_{r\alpha}, \quad \text{where} \quad \omega_{r\alpha} = \begin{cases} 0 & \text{for} \quad x_{r\alpha} = 0\ ,\ E_{r\alpha} < 0 \\ 1 & \text{else} \quad , \end{cases} \tag{20}$$

$$E_{r\alpha} = J_r h_r - \sum_{\beta} I_r^{\alpha\beta} x_{r\beta} + \sum_{r'} \mathsf{f}_{rr'}^{\alpha} x_{r'\alpha}, \tag{21}$$

where $J_r > 0$, $I_r^{\alpha\beta} = I_r^{\beta\alpha} > 0$, and $\mathsf{f}_{rr'}^{\alpha} = \mathsf{f}_{r'r}^{\alpha}$.

Theorem 1. *If $\kappa_{r\alpha} > 0$ with $\kappa_{r\alpha} = I_r^{\alpha\alpha} - \mathsf{f}_{rr}^{\alpha} - \sum_{r' \neq r} \max(0, \mathsf{f}_{rr'}^{\alpha})$, then the CLM dynamics is bounded. If $0 \leq x_{r\alpha}(0) \leq M$ for all r, α, where $M = \max_r (J_r h_r / \kappa_{r\alpha})$ then $0 \leq x_{r\alpha}(t) \leq M$ for all r, α and $t > 0$.*

Proof. We prove the boundedness by constructing a hypercube in the positive domain which cannot be left by the dynamics. Since $\dot{x}_{r\alpha} \geq 0$ for $x_{r\alpha} = 0$ the dynamics cannot leave the positive domain if initialised with $x_{r\alpha}(0) \geq 0$ and therefore $x_{r\alpha}(t) \geq 0$ for all $t \geq 0$. The (r, α)-face of the hypercube is defined by

$$x_{r\alpha} = M, \quad x_{r'\beta} \leq M \quad \text{for} \quad r' \neq r \text{ or } \beta \neq \alpha. \tag{22}$$

We now show how to choose M such that $\dot{x}_{r\alpha} \leq 0$ for all $x_{r\alpha}$ on the face (r, α), causing the system to stay inside the hypercube. On face (r, α) we have $x_{r\alpha} = M > 0$ and therefore $\omega_{r\alpha} = 1$. Hence

$$\dot{x}_{r\alpha} = J_r h_r - \sum_{\beta} I_r^{\alpha\beta} x_{r\beta} + \sum_{r'} \mathsf{f}_{rr'}^{\alpha} x_{r'\alpha} \tag{23}$$

$$\leq J_r h_r - I_r^{\alpha\alpha} x_{r\alpha} + \mathsf{f}_{rr}^{\alpha} x_{r\alpha} + \sum_{r' \neq r} \max(0, \mathsf{f}_{rr'}^{\alpha}) x_{r'\alpha} \tag{24}$$

where we have omitted negative off-diagonal terms. We can now insert (22) and obtain

$$\dot{x}_{r\alpha} \leq J_r h_r - \left(I_r^{\alpha\alpha} - \mathsf{f}_{rr}^{\alpha} - \sum_{r' \neq r} \max(0, \mathsf{f}_{rr'}^{\alpha})\right) M \tag{25}$$

$$= J_r h_r - \kappa_{r\alpha} M \quad \leq 0 \quad \text{if } M \geq J_r h_r / \kappa_{r\alpha} \tag{26}$$

Therefore, if we choose $M \geq \max_r (J_r h_r / \kappa_{r\alpha})$ then $\dot{x}_r \leq 0$ on all faces r. □

Note that the conditions on $\kappa_{r\alpha}$ in Theorem 1 do not imply diagonal dominance of the global interaction matrix $\mathsf{G}_{rr'}^{\alpha\beta}$, since they do not take into account the vertical cross-inhibitions $I_r^{\alpha\beta}$.

Corollary 1. *Under the conditions of Theorem 1 the dynamical system*

$$\dot{x}_{r\alpha} = -x_{r\alpha} + \sigma(E_{r\alpha} + x_{r\alpha}) \tag{27}$$

where $\sigma(x) = \max(0, x)$ is bounded under the same conditions.

Proof. Since $\omega_{r\alpha} E_{r\alpha} \leq 0$ if and only if $-x_{r\alpha} + \sigma(E_{r\alpha} + x_{r\alpha}) \leq 0$, the conditions on the faces of the constructed box hold equivalently for this alternative formulation of the dynamics. □

In the following we state sufficient conditions on the lateral and vertical interactions that ensure the convergence to an unambiguous state, with only one active layer within a column. We define the lateral feedback $F_{r\alpha}$ within a layer by $F_{r\alpha} = \sum_{r'} \mathsf{f}_{rr'}^{\alpha} x_{r'\alpha}$.

Theorem 2. *If the lateral interaction is self-excitatory, $\mathsf{f}_{rr}^{\alpha} > 0$ for all r, α, and the vertical interactions satisfy $I_r^{\alpha\alpha} I_r^{\beta\beta} \leq (I_r^{\alpha\beta})^2$ for all α, β then an attractor of the CLM has in each column r either*

i) at most one positive activity $x_{r\hat{\alpha}}$ with

$$x_{r\hat{\alpha}} = \frac{J_r}{I_r^{\hat{\alpha}\hat{\alpha}}} h_r + \frac{F_{r\hat{\alpha}}}{I_r^{\hat{\alpha}\hat{\alpha}}}, \quad x_{r\beta} = 0 \text{ for all } \beta \neq \hat{\alpha}, \tag{28}$$

where $\hat{\alpha} = \hat{\alpha}(r)$ is the index of the maximally supporting layer characterized by $F_{r\hat{\alpha}} > F_{r\beta}$ for all $\beta \neq \hat{\alpha}$ or

ii) all activities $x_{r\alpha}$, $\alpha = 1, \ldots, L$ in a column r vanish and $F_{r\alpha} \leq -J_r h_r$ for all $\alpha = 1, \ldots, L$.

Proof. Suppose an equilibrium $\mathbf{x}^F$ has two positive activities $x_{r\alpha,\beta}^F > 0$ in a column r at two layers α and β. Then the constraint is inactive, $\omega_{r\alpha,\beta} = 1$, hence $\frac{\partial E}{\partial x_{r\alpha,\beta}} = 0$. Now consider a small perturbation within this column of the form $\epsilon_{r\alpha} = \eta_\alpha$, $\epsilon_{r\beta} = \eta_\beta$, $\epsilon_{r'\alpha'} = 0$ for all other r', α'. Expanding the quadratic function E about the equilibrium we obtain for ΔE under this perturbation with all linear components vanishing at the equilibrium

$$2\Delta E = \sum_{rr'\alpha\beta} (I_r^{\alpha\beta} \delta_{rr'} - \delta_{\alpha\beta} \mathsf{f}_{rr'}^{\alpha}) \epsilon_{r\alpha} \epsilon_{r'\beta} \tag{29}$$

$$= I_r^{\alpha\alpha} \eta_\alpha^2 + I_r^{\beta\beta} \eta_\beta^2 + 2 I_r^{\alpha\beta} \eta_\alpha \eta_\beta - \mathsf{f}_{rr}^{\alpha} \eta_\alpha^2 - \mathsf{f}_{rr}^{\beta} \eta_\beta^2 \tag{30}$$

$$\leq I_r^{\alpha\alpha} \eta_\alpha^2 + I_r^{\beta\beta} \eta_\beta^2 + 2 I_r^{\alpha\beta} \eta_\alpha \eta_\beta \ . \tag{31}$$

If we can state a perturbation $(\eta_\alpha, \eta_\beta)$ for which $\Delta E < 0$ then $\mathbf{x}^F$ can be no local minimum and therefore be no attractor of the gradient-descent dynamics. The quadratic (2×2) form (31) has at least one negative eigenvalue λ_1 or λ_2 if $\lambda_1\lambda_2 = I_r^{\alpha\alpha} I_r^{\beta\beta} - (I_r^{\alpha\beta})^2 < 0$. Then, however, exists a perturbation $(\eta_\alpha, \eta_\beta)$ with $\Delta E < 0$. If equality holds, $I_r^{\alpha\alpha} I_r^{\beta\beta} = (I_r^{\alpha\beta})^2$ then $\lambda_1 = 0$ or λ_2=0 and we can state a perturbation from the null-space for which then $\Delta E < 0$ due to $\mathsf{f}_{rr}^{\alpha}, \mathsf{f}_{rr}^{\beta} > 0$. So, we have shown by contradiction that an attractor can have at most one positive activity in a column.

Now consider case $i)$: Let $\hat{\alpha}$ denote the layer index of the positive activity in row r. The constraint is then inactive for $x_{r\hat{\alpha}}$ and active for $x_{r\beta\neq\hat{\alpha}}$. Hence

$$0 = E_{r\hat{\alpha}} = J_r h_r - I_r^{\hat{\alpha}\hat{\alpha}} x_{r\hat{\alpha}} + F_{r\hat{\alpha}}, \tag{32}$$

$$0 > E_{r\beta\neq\hat{\alpha}} = J_r h_r - I_r^{\hat{\alpha}\hat{\alpha}} x_{r\hat{\alpha}} + F_{r\beta\neq\hat{\alpha}}, \tag{33}$$

which proves case $i)$. Case $ii)$ gives just the remaining possible stable equilibria which are not covered by $i)$. □

Appendix B: Preprocessing and Cocircular Interaction

The edge features are generated by subdividing the image into non-overlapping small subrectangles. On each subrectangle 3x3 pixel Sobel x and y operators are applied which sample the intensity gradient information at orthogonal directions. Within each subrectangle we choose an edge feature indexed by r at the position of maximum squared Sobel response and a unit orientation vector in the direction of the normalized Sobel x and y components. The input h_r is given by the squared Sobel response added by a small positive constant. We observed that this slightly raising of the zero level improves the ability to enhance low input contours through lateral effects. The constant was chosen as ten percent of the maximum edge intensity. The original pixel sizes of the images before subsampling were 298×284 (fruit still life), 224×224 (leena image) and 380×380 (aerial image).

The co-circular interaction of two edges at positions $\mathbf{r}_1 = (r_1^x, r_1^y)$ and $\mathbf{r}_2 = (r_2^x, r_2^y)$ with a difference vector $\mathbf{d} = \mathbf{r}_1 - \mathbf{r}_2$, $d = ||\mathbf{d}||$ and $\hat{\mathbf{d}} = \mathbf{d}/d$, and unit orientation vectors $\hat{\mathbf{n}}_1 = (n_1^x, n_1^y)$, $\hat{\mathbf{n}}_2 = (n_2^x, n_2^y)$ (see Fig.6) is given by

$$\begin{aligned} &f^{\text{cocirc}}((\mathbf{r}_1, \hat{\mathbf{n}}_1), (\mathbf{r}_2, \hat{\mathbf{n}}_2)) \\ &\quad = \theta(a_1 a_2 q)\big(e^{-d^2/R^2 - C^2 S}\big) - I e^{-2d^2/R^2}, \end{aligned}$$

where $a_1 = n_1^x \hat{d}_y - n_1^y \hat{d}_x$, $a_2 = n_2^x \hat{d}_y - n_2^y \hat{d}_x$, $q = \hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2$ and $\theta(x) = 1$ for $x \geq 0$ and $\theta(x) = 0$ otherwise is necessary to exclude skewed symmetric edges. The parameter R controls the spatial range, which is smaller for the inhibitory component. The degree of co-circularity is given by $C = |\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{d}}| - |\hat{\mathbf{n}}_2 \cdot \hat{\mathbf{d}}|$ which is equal to zero if both edges lie tangentially to a common circle. The parameter $S > 0$ controls the sharpness of the co-circularity constraint and $I > 0$ controls the strength of the local inhibition. For all the simulations in this paper, the spatial dimensions were scaled into a unit square, and the parameters were $R = 0.1$, $S = 300$, $I = 0.5$, $m = 3.5$.

References

August, J., Siddiqi, K., & Zucker, S. W. (1999). Contour fragment grouping and shared, simple occluders. Computer Vision and Image Understanding. (to appear).

Baldi, P. & Meir, R. (1990). Computing with arrays of coupled oscillators: An application to preattentive texture discrimination. Neural Computation 2(4), 458–471.

Ben-Yishai, R., Lev Bar-Or, R., & Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. Proc. Nat. Acad. Sci. USA, 92, 3844–3848.

Besag, J., Green, P., Higdon, D., & Mengersen, K. (1995). Bayesian computation and stochastic systems. Statistical Science, 1, 3–66.

Blasdel, G. G. & Fitzpatrick, D. (1984). Physiological organization of layer 4 in macaque striate cortex. J. Neurosci., 4(3), 880–895.

Blatt, M., Wiseman, S., & Domany, E. (1997). Data clustering using a model granular magnet. Neural Computation 9(8), 1805–1842.

Brown, G. J. & Wang, D. L. (1997). Modelling the perceptual segregation of double vowels with a network of neural oscillators. Neur. Netw., 10, 1547–1558.

DeAngelis, G. C., Freeman, R. D., & Ohazawa, I. (1994). Length and width tuning of nerons in the cat's primary visual cortex. Neurophysiology, 71, 347–374.

Douglas, R., Koch, C., Mahowald, M., Martin, K., & Suarez, H. (1995). Recurrent excitation in neocortical circuits. Science, 269, 981–985.

Eckhorn, R. (1994). Oscillatory and non-oscillatory synchronizations in the visual cortex and their roles in associations of visual features. Progress in Brain Research, 102, 405–426.

Elder, J. H. & Zucker, S. W. (1996). Computing contour closure. In Proc. 4th European Conference on Computer Vision, pp. 399–412.

Feng, J. (1997). Lyapunov functions for neural nets with nondifferentiable input-output characteristics. Neural Computation 9(1), 43–49.

Field, D. J., Hayes, A., & Hess, R. F. (1992). Contour integration by the human visual system: Evidence for a local "association field". Vision Research 33(2), 173–193.

Geman, D., Geman, S., Graffigne, C., & Dong, P. (1990). Boundary detection by constrained optimization. IEEE Trans. Pattern Analysis and Machine Intelligence 12(7), 609–628.

Geman, S. & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Trans. Pattern Analysis and Machine Intelligence, 6, 721–741.

Gilbert, C. D. (1992). Horizontal integration and cortical dynamics. Neuron 9(1), 1–13.

Hahnloser, R. L. T. (1998). On the piecewise analysis of linear threshold neurons. Neur. Netw., 11, 691–697.

Hawken, M. J. & Parker, A. J. (1984). Contrast sensitivity and orientation selectivity in lamina iv of the striate cortex of old world monkeys. Exp. Brain Res., 54, 367–372.

Herault, L. & Horaud, R. (1993). Figure-ground discrimination: A combinatorial optimization approach. IEEE Trans. Pattern Analysis and Machine Intelligence 15(9), 899–914.

Hofmann, T. & Buhmann, J. (1997). Pairwise data clustering by deterministic annealing. IEEE Trans. Pattern Analysis and Machine Intelligence 19(1), 1–14.

Hubel, D. H. & Wiesel, T. N. (1962). Receptive fields, binocular interaction adn functional architecture in cat's visual cortex. J. Physiol., 160, 106–154.

Hubel, D. H. & Wiesel, T. N. (1972). Laminar and columnar distribution of geniculocortical fibers in macaque monkey. J. Comp. Neurol., 146, 421–450.

Hummel, R. A. & Zucker, S. W. (1983). On the foundations of relaxation labeling processes. IEEE Trans. Pattern Analysis and Machine Intelligence 5(3), 267–286.

Kamgar-Parsi, B. & Kamgar-Parsi, B. (1990). On problem solving with Hopfield networks. Biol. Cyb., 62, 415–423.

Kapadia, M. K., Ito, M., & Westheimer, G. (1995). Improvement of visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. Neuron, 15, 843–856.

Kimia, B. B., Tannenbaum, A. R., & Zucker, S. W. (1995). Shapes, shocks, and deformations i. the components of 2-dimensional shape and the reaction-diffusion space. Int. Journal of Computer Vision, 15, 189–224.

Kovács, I. & Julesz, B. (1993). A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation. Proc. Nat. Acad. Sci. USA, 90, 7495–7497.

Li, Z. (1998). A neural model of contour integration in the primary visual cortex. Neural Computation, 10, 903–940.

Li, Z. & Dayan, P. (1999). Computational differences between asymmetrical and symmetrical networks. Network, 10, 59–77.

Lund, J. S. (1987). Local circuit neurons of macaque monkey striate cortex: I. neurons of laminae 4c and 5a. Journal of Comparative Neurology, 257, 60–92.

Mumford, D. & Shah, J. (1989). Optimal approximations by piecewise smooth functions, and associated variational problems. Comm. Pure and Appl. Math., 42, 577–684.

Mundel, T., Dimitrov, A., & Cowan, J. D. (1997). Visual cortex circuitry and orientation tuning. In M. C. Mozer, M. I. Jordan, & T. Petsche (Eds.), Advances in Neural Information Processing Systems, Volume 9, pp. 887. The MIT Press.

Ontrup, J. & Ritter, H. (1998). Perceptual grouping in a neural model: Reproducing human texture perception. Technical Report SFB 360 98-6, University of Bielefeld.

Opara, R. & Wörgötter, F. (1998). A fast and robust cluster update algorithm for image segmentaion in spin-lattice models without annealing - visual latencies revisited. Neural Computation 10(6), 1547–1566.

Parent, P. & Zucker, S. W. (1989). Trace inference, curvature consistency, and curve detection. IEEE Trans. Pattern Analysis and Machine Intelligence 46(17), 763–770.

Pelillo, M. (1994). On the dynamics of relaxation labeling processes. In IEEE Intl. Conf. on Neural Networks (ICNN), Volume 2, pp. 606–1294. IEEE Press.

Peterson, C. & Soderberg, B. (1989). A new method for mapping optimization problems onto neural networks. Int. J. Neural Systems 1(1), 3–22.

Polat, U. & Sagi, D. (1994). The architecture of perceptual spatial interactions. Vision Research, 34, 73–78.

Ritter, H. (1990). A spatial approach to feature linking. In Proc. Int. Neur. Netw. Conf. Paris Vol 2., pp. 898–901.

Rockland, K. S. & Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. J. Comp. Neurol., 216, 303–318.

Rose, K. & Fox, G. (1993). Constrained clustering as an optimization method. IEEE Trans. Pattern Analysis and Machine Intelligence 15(8), 785–794.

Rosenfeld, A., Hummel, R. A., & Zucker, S. W. (1976). Scene labeling by relaxation operations. IEEE Trans. Systems, Man, Cybernetics 6(6), 420–433.

Sajda, P. & Finkel, L. H. (1994). Dual mechanisms for neural binding and segmentation. In J. D. Cowan, G. Tesauro, & J. Alspector (Eds.), Advances in Neural Information Processing Systems, Volume 6, pp. 993–1000. Morgan Kaufmann Publishers, Inc.

Schillen, T. B. & König, P. (1994). Binding by temporal structure in multiple feature domains of an oscillatory network. Biol. Cyb., 70, 397–405.

Shi, J. & Malik, J. (1997). Normalized cuts and image segmentation. In Proc. of the IEEE Conf. on Comp. Vision and Pattern Recognition, Puerto Rico.

Singer, W. & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. Ann. Rev. Neurosci., 18, 555–586.

Somers, D. & Kopell, N. (1993). Rapid synchronization through fast threshold modulation. Biol. Cyb., 68, 393–407.

Sompolinsky, H., Golomb, D., & Kleinfeld, D. (1991). Cooperative dynamics in visual processing. Physical Review A 43(12), 6990–7011.

Sum, J. P. F. & Tam, P. K. S. (1996). Note on the maxnet dynamics. Neural Computation 8(3), 491–499.

Terman, D. & Wang, D. L. (1995). Global competition and local cooperation in a network of coupled oscillators. Physica D, 81, 148–176.

von der Malsburg, C. (1981). The correlation theory of brain function. Technical Report 81-2, MPI Göttingen.

von der Malsburg, C. (1995). Binding in models of perception and brain function. Current Opinion in Neurobiology, 5, 520–526.

von der Malsburg, C. & Buhmann, J. (1992). Sensory segmentation with coupled oscillators. Biol. Cyb., 54, 29–40.

Wang, D. & Terman, D. (1997). Image segmentation based on oscillatory correlation. Neural Computation 9(4), 805–836.

Weliky, M., Kandler, K., Fitzpatrick, D., & Katz, L. C. (1995). Patterns of excitation and inhibition evoked by horizontal connections in visual cortex share a common relationship to orientation columns. Neuron, 15, 541–552.

Wersing, H. & Ritter, H. (1999). Feature binding and relaxation labeling with the competitive layer model. In Proc. Eur. Symp. on Art. Neur. Netw. ESANN, pp. 295–300.

Wersing, H., Steil, J. J., & Ritter, H. (1997). A layered recurrent neural network for feature grouping. In W. Gerstner, A. Germond, M. Hasler, & J.-D. Nicoud (Eds.), Proc. ICANN, pp. 439–444.

Yabuta, N. H. & Callaway, E. M. (1998). Functional streams and local connections of layer 4c neurons in primary visual cortex of the macaque monkey. J. Neurosci., 18(20), 9489–9499.

Yen, S.-C. & Finkel, L. (1996). Salient contour extraction by temporal binding in a cortically-based network. In M. Jordan, M. Mozer, & T. Petsche (Eds.), Advances in Neural Information Processing Systems 9, pp. 915–921. MIT Press.

Yen, S.-C. & Finkel, L. H. (1998). Extraction of perceptually salient contours by striate cortical networks. Vision Research 38(5), 719–741.

Yoshioka, T., Levitt, J. B., & Lund, J. S. (1994). Independence and merger of thalamocortical channels within macaque monkey primary visual cortex: anatomy of interlaminar projections. Vis. Neurosci., 11, 467–489.

Zucker, S. W., Dobbins, A., & Iverson, L. (1989). Two stages of curve detection suggest two styles of visual computation. Neural Computation 1(1), 68–81.