COMPUTATIONAL NEURAL MODELS OF SPATIAL INTEGRATION IN PERCEPTUAL GROUPING

Heiko Neumann† and Ennio Mingolla‡
† Fakultät für Informatik, Abteilung Neuroinformatik, Universität Ulm,
‡ Department of Cognitive and Neural Systems, Boston University

ABSTRACT

Recent developments in the neural computational modeling of perceptual grouping are described with reference to a newly proposed taxonomy to formalize mechanisms of spatial integration. This notational framework and nomenclature is introduced in order to clarify key properties common to all or most models, while permitting unique attributes of each approach to be independently examined. The strength of spatial integration in the models that are considered is always some function of the distances and relative alignments in perceptual space of the centers of units representing orientational features or energy in a visual scene. We discuss the significance of variations of the constituents of an activation function for spatial integration, and also consider the larger modeling framework in which this function is applied in each approach. We also discuss the relationship of feedforward and feedback mechanisms and the issues of self-organization as core principles underlying the establishment of spatial integration mechanisms. The relationship of the grouping models to models of other visual competencies is considered with respect to prospects for future research.
INTRODUCTION

Recent years have seen burgeoning interest in spatial integration mechanisms for perceptual grouping, with psychophysical, physiological, brain imaging, and computational perspectives all playing important roles. This chapter examines formal approaches to a particular aspect of grouping related to the Gestalt concept of good continuation (Koffka, 1935) that have garnered increasing attention since Field, Hayes, and Hess (1993) popularized the notion of an “association field,” a figure-eight shaped zone aligned with perceptual contour segments, along which facilitatory perceptual interactions with other contour segments tend to occur. We examine prior expositions of the figure-eight idea and more recent developments, with the focus on studies that have included sufficient formal specification to afford computer implementation and simulation of the key grouping mechanism and its ancillary machinery, such as early filters.

Our focus is on models intended to address psychophysical and physiological data, as opposed to pure computer vision approaches. The key behavioral properties of spatial integration include enhancement or suppression of detectability of contour segments depending on particular geometric relationships among local orientations and positions of contour segments to be grouped, as are described in Section 2. Important constraints from neural data include the results of pioneering studies of von der Heydt, Gilbert, and their collaborators concerning the physiology and anatomy of cells with long-range horizontal connections in cortical areas V1 and V2.

Even with such restrictions on scope, a number of models remain to be examined, and our classification is based on several factors. These include the type of formalism employed, such as relaxation labeling or systems of ordinary differential equations. We also compare functional properties of models, particularly with respect to their treatment of feedback, whether in the form of recurrent horizontal connections among network nodes, or in the form of processing loops involving nodes of different layers. The common thread in the models we address is that some key stage or process in the model accomplishes spatial integration using some sort of figure-eight based mechanism or process. We next make precise the significance of this construct, before examining several models’ treatment of it in detail, by describing each model’s formalization of spatial integration and presenting some key simulation results. A concluding discussion outlines outstanding questions for further modeling studies.
Figure 1. The “bipole icon” for modeling feature integration in spatial grouping (top). The quantity to be assessed is the “contribution” of activation in a oriented unit, denoted by the light ellipse, to the activation of another unit, denoted by the dark ellipse. The figure-eight shape of a bipole expresses relations among three fundamental quantities (bottom): (a) the distance, $r$, between the centers of the two ellipses, (b) the angle, $\vartheta$, formed by a ray passing through the centers of the two ellipses and the orientation axis of the dark ellipse, and (c) the difference in orientation, $\varphi$, of the principle axes of the two ellipses.

MECHANISMS OF SPATIAL INTEGRATION

The bipole icon

The geometry of spatial integration is summarized in the bipole\textsuperscript{1} icon of Figure 1. Imagine that some pattern of contrast, such as an edge, occurs in a scene in the region of, and in the orientation denoted by the long axis of, the light ellipse (at location $x'$). How should the influence of that edge on the representation of contour strength in the region of the dark ellipse be calculated? Whatever the details of the computation, the fundamental quantities to be considered are: (1) the distance, $r$, between the centers of the two ellipses; (2) the angle, $\vartheta$, formed by a ray passing through the centers of the two ellipses and a ray along the principle axis of the dark ellipse; and (3) the difference in orientation, $\varphi$, of the principle axes of the two ellipses. The characteristic “figure-eight”

\textsuperscript{1}The term “bipole” was introduced by Grossberg and Mingolla (1987) to describe a unit in a neural network model described by Cohen and Grossberg (1984b), Grossberg (1984), and Grossberg and Mingolla (1985a, 1985b) as a “cooperative cell”.
Figure 2. Left panel: This plot summarizes psychophysical data from 450 trials of one observer in a tilt illusion experiment. The observer controlled the orientation of a central line segment displayed in juxtaposition with two symmetrically flanking lines whose actual orientation was 5 degrees from the vertical. The observer's task was to make the central line appear vertical. The data show regions for which attractive (light) and repulsive (dark) tilt illusions were obtained. Right panel: Summary of excitatory and inhibitory zones of influence for flanking stimuli on single cells of vertical orientational preference whose receptive fields are centered in the diagram. See Kapadia et al. (2000) for details. (This is Figure 12a of Kapadia et al. (2000), reproduced with permission. Efforts to convert the original figure from color to grayscale may have introduced unintended artifacts. Although the overall bipole shape of excitatory regions is representative, readers should consult the original source if possible.)

The bipole icon can be interpreted in several rubrics. Most generally, it expresses the basic functional task that must be achieved in contour integration. That task is to assess the influence that a contour segment at one location should have on the likelihood of a contour being perceived at a second location — specifically, at a place “pointed to by” or “consistent with” the first segment. Instead of likelihood, we might also speak of the contour at the former location modulating the perceived strength or salience of a contour at another. In this respect, the icon summarizes functional relationships in the processing of visual input by our perceptual system. Before examining anatomical and physiological correlates of the bipole shape, we first turn to psychophysical evidence.
Psychophysical and physiological data

Recent investigations by Kapadia et al. (2000) arrived at a dipole shape in a summary data plot for an experiment in which observers were asked to adjust the orientation of a test line of dimension 8 x 1 minutes in the presence of flanking lines. Figure 2 summarizes regions for which flanking lines placed in individual grid squares affect the perceived orientation of a line in the central square, with light regions coding zones of attraction and dark regions coding zones of repulsion in the perceived orientation of a vertical line. The height of each square corresponds to the arc length of a test or flanking line (8 minutes).

That the region of orientational attraction in the left panel of Figure 2 takes the shape of a dipole is both a hint about the human visual system’s perceptual processing and a fact in need of explanation. One possible hypothesis is that our visual systems contain neurons at some layer whose pattern of connections with neurons that synapse upon them from earlier layers is itself that of a dipole. Note that this is not a logical necessity. That is, the pattern of psychophysical data of Figure 2 could be the result of processing by the subjects’ entire visual systems, no individual component of which need have the structure of a dipole. The final pattern of data could instead be the result of many interacting components of unknown character. That is, the receptive field of a cell is the result of all factors that bear on a cell’s potential, including bottom-up, top-down, and lateral excitatory and inhibitory connections. (See Figure 3.) Kapadia et al. (2000) made progress on this question by performing electrophysiological recordings of activity of units in V1 of macaques, using stimuli analogous to those used in the psychophysical task performed by human subjects. The zones of attraction and repulsion in the single-unit studies bear a striking resemblance to those found psychophysically. (See Figure 2.)
Figure 4. Summary of specificity of horizontal connections between V1 cells of layer 2/3. Axon arboreizations from two example cells are shown over a combined map of visual space and orientation preference (dark regions: cells with 90 deg. orientation preference; light regions: cells with 0 deg. preference). Horizontal connections are reciprocal such that layer 2/3 cells receive input from other cells in layer 2/3 with the same orientation selectivity. The cell shown above that was found in a dark region of the map projects to other areas with the same orientation preference (90 deg.). Moreover, those cells that were selectively contacted lie along a line that is orthogonal to the 0 deg. meridian. The other cell shown was found in a light region of the map and projects to spatially distant cells with the same orientation preference (0 deg. in this case) and that lie along a line parallel to the 0 deg. meridian. (This is Figure 12 of Bosking et al. (1997), reproduced with permission.)

Anatomical data

Recent anatomical data support the “bipole hypothesis” – the claim that the figure-eight shaped region of excitatory interactions in perceptual contour integration is mediated primarily by direct anatomical linkages that form a figure-eight shape in retinotopic coordinates (Grossberg and Mingolla, 1985a, 1985b, 1987). Early work had already shown that horizontal connections among single cells in V1 connect cells of like orientational preference (Gilbert and Wiesel, 1979, 1983; Martin and Whitteridge, 1984; Rockland and Lund, 1982). More recently Bosking et al. (1997) have used a double-labeling technique to demonstrate directly that such connections in tree shrew are also preferentially made among cells whose receptive field centers in retinotopic space are in positions aligned with the orientational preference of the cell making the long-range connections. That is, a vertically tuned cell will make long-range connections with cells that are also vertically tuned and that code positions above or below, but not to the side of, its receptive field center. (See Figure 4.)
Figure 5. Feedforward input convergence onto a target cell $v_t$ over a cascade of neural layers. The arrows connecting the top two rows of open circles stand for pathways to provide input for the target cell. These pathways may have different efficacies and can be modeled as kernels of different weight coefficients. The kernels thus define the structure of direct connections to cell $v_t$. The squares at the “input layer” (e.g., retina) code the receptive field (RF) of $v_t$. The RF of a cell thus determines, functionally, where in the periphery a stimulation will yield a response at the cell. The RF is always defined in the coordinates of the peripheral sensor, e.g., the retina.

Bipoles in modeling

Before proceeding, then, it is useful to review some points concerning usage of terms in computational modeling, psychophysics, and physiology. A physiologist may use electrode recordings to determine a cell’s receptive field. Recently, there has been an explosion of interest cells with so-called “nonclassical receptive field” properties—that is, cells whose firing can be modulated by stimulus manipulations well outside the classical receptive field. (See Spillmann and Werner (1996) for a review.) Note, however, that whether classical or nonclassical, a cell’s receptive field is, by definition, a functional property—the result of all interactions from receptors through any “top-down” feedback that might bear on a cell’s potential. The present paper, on the other hand, examines computational variations of a hypothesis concerning how a specific bipoles pattern of connectivity exists among relatively tightly-coupled neural units. This pattern of connectivity is described in formal terms by a kernel, or set of weights, coding the strength of connection between a spatial array of units in a network to a particular designated unit. (See Figure 5.)

The direct analogue of a kernel in vivo, however, is difficult or impossible to assess for most cortical neurons. To precisely determine all the cells that send input to a particular neuron, and with what “weight,” would be an heroic undertaking. Note that Figure 4 shows the pattern of axonal connections radiating outward from a single neuron, which is
Figure 6. Taxonomy of contributions to the formal definition of spatial integration mechanisms. Their functionality can be organized in a tree structure of elements. The spatial weighting defines the structure of lateral spatial connectivity, or kernel. The feature compatibility function defines the representation of relatable visual structure, or context. This encodes the model of the expected most likely visual arrangement in the environment. The product of spatial weighting and compatibility function defines the total coupling strength for support measures in the space-feature domain. The definition of individual components are described in the text.

the dual of the set of connections terminating at a neuron (cf. Grossberg and Mingolla, 1985b, Figures 32 and 33). In any event, since the pioneering work of von der Heydt, Peterhans, and Baumgartner (1984) first provided evidence for bipole shaped receptive fields, the idea that a structural bipole kernel underlies the functional receptive field properties has featured in many models. We next explore the variations employed in detailed instantiations of such kernels in several key modeling studies.

ELEMENTS OF SPATIAL INTEGRATION FOR GROUPING

Dimensions of integration

Grouping mechanisms invoke principles of selective integration of more localized measurements along any of several possible feature dimensions. The primary goal of this section is to describe the elements of spatial integration mechanisms and their formal description. Figure 6 depicts a taxonomy of the elements which contribute to the definition of the support and final activation function computed as a mechanism of feature integration and grouping.
In all, the result of integration of items defines the support of a localized target feature measurement at a given spatial location based on the configuration of other measurements. Based on the bipole structure of left and right sub-fields, the support can be defined as a function

$$\text{support}_{xy, \text{feature}} = \{\text{left-input}_{xy, \text{feature}}\} \circ \{\text{right-input}_{xy, \text{feature}}\}$$

where ‘$\circ$’ denotes some operation to define the combination of sub-fields. Subscripts $xy$ denote spatial locations in a 2-D retinotopic map and $\text{feature}$ identifies the feature dimension involved in the grouping process. In the bipole model the segregated lobes of the integration field define the support of a target item. The activation of the corresponding target cell results as a function of the support such that

$$\text{activation}_{xy, \text{feature}} = f(\text{support}_{xy, \text{feature}}).$$

The formal description of the mechanisms underlying the computation of activations ‘left-input’ and ‘right-input’, respectively, necessitates the detailed specification of the interaction of activities in grouping.

A spatial weighting, or kernel, function specifies the mutual influence between two localized features. Items that are closely spaced are in general more likely to be grouped than candidates that are located far apart. The underlying spatial neighborhood function often selectively facilitates a sector of a spatial surround to define an anisotropic coupling that is compatible with the feature domain. Elementary features along dimensions such as, e.g., (tangential) orientation, motion direction, and disparity provide the dimensions of the visual representation space. The feature relatability (or compatibility; see Parent and Zucker (1989) and Kellman and Shipley (1991)) between stimulus items is defined along these dimensions. Herein, the most typical or likely appearance of meaningful structure is somehow encoded to represent “what feature goes with what”. The resulting model representation is encoded in the spatial weights of the connectivity pattern of a target cell with its surround and thus defines a spatial filter (cf. Grossberg, 1980). In the following treatment, we will particularly focus on grouping mechanisms for static shape or form processing and, therefore, consider only orientation as the relevant feature dimension. (See Figure 1.)

In most models the connectivity pattern of the relatable features is pre-specified, or programmed, based on a framework of geometric measures. These are encoded in closed-form mechanisms that define a static efficacy between tuples, e.g. pairs, of feature measurements at given locations. In this sense, they are designed a priori in order to achieve a given functionality. To date, few approaches investigate the possible self-organization of such lateral interactions in a neural architecture. Note that the spatial
weighting function and the feature relatability define the components of the net coupling strength. This separable function specifies a metric for the similarity measure in the \(<xy, feature>-space and thus defines the distance function for the clustering of a visual pattern according to a relatability measure. This measure contributes to the definition of feature cooperation. In order to get activated, input activations are necessary to gather any support. Here, two different types of interactions can be distinguished: (1) a convergent feedforward mechanism is defined when the bottom-up input is integrated at the target location, whereas (2) a mechanism of (non-linear) lateral interaction is defined when activity is horizontally integrated within a neural layer (compare Figures 3a and 3b).

Taken together, the support — which corresponds to the activity derived from the feature integration process using the bipole mechanism — is usually computed by the following scheme assembled from the elements that were introduced above:

$$\text{support}_{x\theta} = \sum_{x'\phi} \left\{ \text{act}_{x'\phi} \cdot \text{relate}_{\theta\phi} \cdot \text{weight}_{xx\theta}^{\text{left}} \right\} \circ \sum_{x'\phi} \left\{ \text{act}_{x'\phi} \cdot \text{relate}_{\theta\phi} \cdot \text{weight}_{xx\theta}^{\text{right}} \right\} \quad (3)$$

where ‘weight’ denotes the spatial weighting kernel, ‘relate’ the feature relatability, and ‘act’ the input activations (Figure 6). The small circle denotes some nonlinear combination of terms. Coordinates in the space-feature domain are denoted by bold Latin letters (spatial) and by Greek letters (orientation features). Here, \(x = (x, y)\) and \(\theta\) correspond to the location and orientation of the target feature, respectively. Other parameters refer to the specific location \(<x', \phi>\) in the space-orientation neighborhood. (See Figure 1.)\(^2\) The precise definition of these components varies for different models.

**Outline of core models and their components**

Several computational models for feature integration and grouping have been published. We here focus on those approaches which we believe are most influential with respect to (1) the explanation of empirical data and (2) their importance for subsequent scientific developments. In particular, we outline the Boundary Contour System (BCS) model developed by Grossberg and colleagues, the relaxation approach for feature compatibility labeling developed by Zucker and colleagues, and the filter model developed by von der Heydt, Peterhans and colleagues based on their results of neurophysiolog-

\(^2\)Note that in these definitions the feature orientations \(\theta\) and \(\phi\) are measured against an oriented coordinate axis. In Figure 1 a gauge coordinate system is assumed such that the axis of orientation of the target cell serves as the reference. For completeness, the angles which define the relations between quantities can be computed as \(\vartheta = \tan^{-1}(y'/y, x' - x) - \theta\) and \(\varphi = \phi - \theta\).
Neural Models of Spatial Integration

New contributions of other recent models are then introduced and discussed.

For clarity we first establish a coherent usage of variables and names for different types of elements in each model. The following table shows the three generic categories for the definition of the components in different schemes of feature integration. The elementary categories in all models are weighting functions, activations and parameters (constants) of the activation dynamics. Further necessary elements in the formal description will be specified in addition.

<table>
<thead>
<tr>
<th>weighting functions</th>
<th>activations</th>
<th>parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>oriented kernels (bipole)</td>
<td>isotropic kernels</td>
<td></td>
</tr>
<tr>
<td>weight&lt;sub&gt;left&lt;/sub&gt;, weight&lt;sub&gt;right&lt;/sub&gt;</td>
<td>Λ (space), Ψ (orientation)</td>
<td>c, u, v, w, x, y, z, α, β, ...</td>
</tr>
</tbody>
</table>

Boundary Contour System (BCS)

The following section follows the description of the BCS given by Mingolla, Ross, and Grossberg (1999). The BCS consists of a series of boundary detection, competition, and cooperation stages as shown in the block diagram in Figure 7. Stage 1 models the contrast enhancement resulting from on-center, off-surround (ON channel) and off-center, on-surround (OFF channel) interactions at the retina and LGN. These ON and OFF cells compensate for variable illumination by computing locally normalized contrast ratios throughout the image.

At Stage 2, these ON and OFF cells generate half-wave rectified outputs which together drive the activation of oriented simple cells. Simple cells compute a measure of the local image gradient magnitude and orientation. Like-oriented simple cells sensitive to opposite contrast polarities or directions-of-contrast pool their activations at complex cells. Complex cells are hereby rendered insensitive to direction-of-contrast (dark-to-light vs. light-to-dark), as are all subsequent BCS processing stages.

Next, complex cell activations compete at Stage 3 processing. Competition occurs through on-center off-surround processing across both image space (spatial competition) and across orientation space (orientational competition). Spatial and orientational competition captures the functional implications of lateral inhibition across a cortical map in which nearby cells tend to be sensitive both to contrasts at neighboring (or overlapping) image locations and similar boundary orientations. Functionally, competition sharpens boundary localization and orientational tuning of individual complex cell filters. It also
endows the complex cells with a property of endstopping that enables them to respond more vigorously near the ends of a line than at its middle. This competition is also driven by feedback from Stage 4 long-range boundary cooperation, thereby suppressing weaker boundary activations while enhancing the contrast and the completion of stronger and more globally consistent boundary activations.

Long-range boundary cooperation at Stage 4 accomplishes the grouping together of consistent boundaries and the completion of camouflaged boundaries. This cooperation is achieved by bipole cells which realize a type of statistical AND gate, since they fire if both halves of their receptive fields are sufficiently activated by appropriately oriented input contrasts from the complex cells of Stage 3.

The cooperative-competitive (CC) feedback loop between Stage 3 and Stage 4 acts to complete and enhance spatially and orientationally consistent boundary groupings while inhibiting inconsistent ones. This feedback process simultaneously achieves the addition of sharp completions and the suppression of noise. Furthermore, excitatory and inhibitory feedback are balanced such that boundary strength sensitivity is preserved and boundary activations are not driven to saturation. The robust preservation of sensitivity to the analog strength of inputs that support long-range completion of boundary signals over gaps in the image — that is, regions of the image where local image data would not signal a boundary — is an important innovation of recent versions of the BCS (Grossberg et al., 1997; Mingolla et al. 1999; Ross et al. 2000). Previous versions of the BCS tended to sacrifice sensitivity to input strength as a cost of the advantages of context-sensitive long-range grouping in a feedback system. The present architecture combines the nonlinear choice properties necessary to determine whether and where to coherently link boundary segments with sensitivity to input contrast magnitudes in the completed boundary signals.

The BCS boundaries act as barriers to diffusion of the ON and OFF contrast signals within the FCS. Cohen and Grossberg (1984a), Grossberg and Todorović (1988), Neumann (1996), Pessoa, Neumann, and Mingolla (1995), and Neumann, Pessoa, and Mingolla (1998) developed the FCS model to simulate data about human brightness perception. The combination of BCS boundary completion and FCS surface diffusion mechanisms is an early exemplar of a class of image processing procedures known as inhomogeneous or geometry-driven diffusion. Unlike most such approaches, however, BCS/FCS algorithms do not require that diffusion rates be limited by initial image data, or by iterative updates of any transformations of image data that are influenced by diffusion. Instead, the boundary signals that limit diffusion are computed by a self-equilibrating process that is both buffered from the effects of the diffusion process and capable of generating barriers to diffusion that are based on contextual image data at
some distance from a pixel location, rather than simply on local measures of image gradients.

**Figure 7.** The left panel shows the macrocircuit of the BCS-FCS model. The right panel shows details of the stage of oriented contrast detection (OC Filter) and of the stages of the competitive-cooperative (CC) Loop of recurrent boundary processing in the BCS. See text for details.

**Description of the modeling framework.** The key elements in the definition of the spatial integration mechanisms of the BCS proposed by Grossberg and Mingolla (1985b) are summarized below. Since several instances of the precise definition of components have been published, we simply describe one selected scheme and mention the alterations.

1. **Spatial weighting functions (bipoles)**

   The weighting functions for the spatial integration are defined by two separate collinearly oriented kernels for the gauge orientation $\theta$. These kernels are defined by a polar separable function denoted by

   \[
   \text{weight}^{\text{left}}_{XX'\theta} = \left[W_{XX'}^{rad} \cdot W_{XX'\theta}^{ang}\right]^+ \quad \text{and} \quad \text{weight}^{\text{right}}_{XX'\theta} = \left[-W_{XX'}^{rad} \cdot W_{XX'\theta}^{ang}\right]^+,
   \]

   respectively, with $[x]^+ = \max[x, 0]$. The radial and angular distance weightings are defined by

   \[
   W_{XX'}^{rad} = \text{elongated-Gaussian}(\mu, \text{dist}), \quad \text{with} \quad \text{dist} = \left((x - x')^2 + (y - y')^2\right)^{1/2},
   \]
\[ W_{xx}^{\text{ang}} \theta = \pm \cos^T(Q_{xx'} - \theta), \text{ with } Q_{xx'} = \tan^{-1}((y' - y)/(x' - x)). \] (5)

The parameter \( \mu \) in the radial weighting denotes the spatial offset of the Gaussian, \( T \). The sign changes distinguish the left and right lobes of the dipole sub-fields. (See Figure 7.)

(2) \textit{Feature cooperation}

The \textit{relatability} of two segregated features in orientation space is encoded in a weighting pattern that represents a model of the most likely underlying geometrical shape that occurs in the visual world. This representation has been sampled at discrete locations. Maximal relatability is defined by orientations tangent to the model segment at the sample locations. Most often shape segments such as straight lines, circular arcs, splines, or parabolic curves have been used. The simplest model utilizes the deviation from the orientation of the virtual line between sample points and the orientation in the neighborhood. More formally, we get

\[ \text{relate}^{\text{straight}}_{xx} \theta \phi = \cos(R(Q_{xx'} - \phi)), \text{ with } Q_{xx'}, \] (6)

with \( R \) denoting a tuning constant (\textit{cf.} Grossberg and Mingolla, 1985b). For a circular shape model the relatability can be formalized by the \( \infty \)-circularity constraint

\[ \text{relate}^{\text{cire}}_{xx} \theta \phi = \cos(R(2Q_{xx'} - (\theta + \phi))), \] (7)

with \( R \) again denoting the tuning of the selectivity (\textit{cf.} Gove \textit{et al.}, 1995; Parent and Zucker, 1989; and several later models).\(^3\)

The \textit{input activation} is represented by a dipole field composed of antagonistic ON and OFF responses in an orientation field, whereby a horizontally tuned ON cell at one location is equivalent to a vertical OFF, and vice versa, and similarly for other pairs of perpendicular orientations. Pairs of cells with mutually orthogonal orientation preferences incorporate a push-pull mechanism for fast reset processes. Their activity (from the previous neural layer) is denoted by

\[ y_x \theta; \quad \overline{y}_x \theta = y_x \theta_\perp. \] (8)

Such pairs of neurons enter in a mutually inhibitory fashion into the integration mechanism.

\(^3\)Note that in order to fulfill the relatability constraint of Kellman and Shipley (1991), additional constraints must be taken into account, namely \( \theta < Q_{xx} < \phi \) (or instead \( \theta > Q_{xx} > \phi \)) and \( |\theta - \phi| \leq 90^\circ \).
(3) **Support and activation function** The support is computed by the combination of sub-fields and their result from integrated activities. We get

\[
\begin{align*}
\tilde{z}_{x\theta} &= g \left( \sum_{x'\phi} \left[ y_{x'\phi} - \bar{y}_{x'\phi} \right]^+ \cdot \text{relate}_{xx'\theta\phi} \cdot \text{weight}_{xx'\theta}^{\text{left}} \right) + \\
&\quad g \left( \sum_{x'\phi} \left[ y_{x'\phi} - \bar{y}_{x'\phi} \right]^+ \cdot \text{relate}_{xx'\theta\phi} \cdot \text{weight}_{xx'\theta}^{\text{right}} \right),
\end{align*}
\]

with \( g(s) \approx s/(\alpha + s) \). When \( \alpha \to 0 \), this function approximates a step or switch and thus allows for setting a threshold such that the target cell responds only if it receives input from both lobes. As a consequence, the mechanism acts like an AND gate, despite additively combining the lobes of a bipole.

The activation of a bipole cell enters a stage of center-surround feedback competition which in turn completes the closed-loop computation of the competitive-cooperative (CC) mechanism. This leads to the dynamics

\[
\begin{align*}
v_{x\theta} &= \frac{h(z_{x\theta})}{1 + \sum_{x'} h(z_{x\theta}) \cdot \Lambda_{xx'}^{FF}}, \quad \text{with} \quad h(s) = \max[s - L, 0], \quad \text{and} \quad (10) \\
\frac{d}{dt} w_{x\theta} &= -w_{x\theta} + I + x_{x\theta} + v_{x\theta} - w_{x\theta} \sum_{x'} x_{x'\theta} \cdot \Lambda_{xx'}^{FF}, \tag{11}
\end{align*}
\]

with \( L \) denoting a threshold firing level, \( x_{x\theta} \) denoting the initial oriented contrast activity from cortical complex cells, and \( I \) is a tonic input. The kernels \( \Lambda \) denote the weighting functions for the spatial competitive interactions in the feedback (\( FB \)) and feedforward (\( FF \)) pathways. The distributions of \( w \)-activity are then transformed via a stage of dipole competition to generate responses \( y_{x\theta} \).

This network incorporates additive feedback processing and thus can generate activities at locations that lack any initial driving input from the input layers. Thus, according to the classification of Crick and Koch (1998), the recurrent network incorporates a **strong loop**. Like-oriented flanking items of proper spacing enable the bipole mechanism to generate an inward directed completion of the activity distribution and thus create illusory contours. Figure 8 shows the action of the spatial completion process.

**Characteristic simulations.** Figure 9 illustrates how the property of analog sensitivity allows the completion of illusory contours by BCS feedback circuits to vary in strength as a function of the “support ratio” of inducers, which is the length of an illusory contour compared to the total (real plus illusory) contour length (Shipley and Kellman, 1992). Note that illusory contour strength increases as an inverted-U shaped function.
From Fragments to Objects

Figure 8. Principle of the feedback mechanism to generate a spatially continuous arrangement of activity in boundary completion processes. Initially, like-oriented contrast cells feed the lobes of a bipole cell at the integration stage (1); activity is fed back to the input layer creating an activation in the middle of the gap (2); these activities in turn are integrated by other bipole cells (3) which in turn also generate new activities (4). This process of rapid completion is finished after two cycles of iteration generating a spatially continuous boundary representation.

of line density, as reported by Lesher and Mingolla (1993). These results were predicted by an analysis of the interaction of short-range inhibition and long-range excitation in the CC Loop. While the addition of arcs to a display containing few inducers initially increases illusory contour strength, inhibition reduces the effective input to the BCS’s bipole completion mechanisms as line ends become too crowded.

Figure 10 illustrates how the bipole spatial integration mechanisms of the BCS can improve performance of image processing algorithms designed to enhance the interpretability by humans of images derived from artificial sensors (Mingolla et al., 1999). Pixel intensities in synthetic aperture radar images span five orders of magnitude, with a relatively few high intensity pixels and a large majority of low intensity ones. Typically, a compressive logarithmic transformation of intensities is used simply to display image data (e.g., on an 8-bit grayscale). Form-sensitive processing by algorithms based on the BCS and FCS models results in imagery with reduced graininess, higher contrast, and greater ease of use than conventional alternatives. The comparison of Figure 10e and Figure 10f, in particular, illustrates the advantages of using bipole completion mechanisms in the CC Loop to enhance favored boundaries, while suppressing less coherent bottom-up boundary signals caused by noise.

Recent work on the BCS model. The original BCS has been extended and elaborated in recent years, as part of a larger model of earlier vision known as the FACADE (Form And Color And DEpth) model (Grossberg, 1994). An important theme of this work has been the identification of specific neuron types in the laminae of cortical ar-
Figure 9. Model simulations of psychophysical data: (Left) In response to the edge inducers, illusory contour strength increases with support ratio. Support ratio is the ratio of real to total contour length. (Right) For the line end inducers, contour strength is an inverted U function of the number and density of line end inducers. Contour strength was determined by computing the average cell activity along the path of the illusory portion of the contour. (Reprinted with permission from Grossberg et al., 1997.)

areas V1 and V2 that instantiate the bipole property (Grossberg, 1999; Grossberg et al., 1997; Ross et al., 2000). Since long-range lateral connections in cortex are axonal, a key question underlying this research is how inward completion, or interpolation without extrapolation, can be achieved in a circuit whose main long-range connections are outward. This work has also shown how analog-sensitive completion can be instantiated in a feedback network. Analog sensitivity refers to the modulation of the perceptual “strength” of completion as a function of variations of contrast, distance, or orientation of inducing elements.

Additional work on the bipole property has shown how a model of laminar mechanisms in areas V1 and V2 can explain difficult recent psychophysical and physiological data concerning how variations of inducer contrast can produce facilitatory or inhibitory effects on single cell activity, while subserving perceptual completion (Grossberg, 1999; Grossberg and Raizada, 2000). These same circuits are also shown to be involved in attentional modulation of activity in early cortical areas such as V1 or V2.

Other work has developed algorithms based on bipole mechanisms and BCS/FCS interactions for processing of synthetic aperture radar images (Mingolla et al., 1999).
Figure 10. (a) Unprocessed synthetic aperture radar (SAR) image of an upstate New York scene consisting of a highway with a bridge overpass. (b) Logarithm-transformed SAR image. (c) Normalized contrast enhancement of (a), output of Stage 1 of BCS of Figure 7a. (d) BCS boundaries, output of Stages 3 and 4 of BCS. (e) Filling-in of featural signals from Stage 1, using Stage 2 boundaries only. (f) Filling-in of featural signals from Stage 2, using bipole-completed boundaries from Stages 3 and 4. See text for details. (Reprinted with permission from Mingolla et al. (1999).)
Another study has shown how dipole mechanisms and related receptive field properties can be self-organized through developmental tuning (i.e., learning) in networks exposed to visual inputs containing image gradients (Grossberg and Williamson, 2001). Dipole mechanisms also underlie a treatment of T-junctions that supports amodal and modal completion and figure/ground separation (Kelly and Grossberg, 2000).

**Relaxation labeling for contour finding**

Relaxation labeling schemes have been developed in the context of optimization to find consistent interpretations for measurement problems with uncertainty. To introduce concepts, consider a set of entities and the relations between them. This can be described as a graph \( G(\mathcal{N}, \mathcal{E}) \) where the set of nodes (or cells), \( \mathcal{N} \), represent the entities (e.g., objects or features) and the links, \( \mathcal{E} \), represent the relations between pairs of nodes. To each of the nodes will be associated a label from a given set, \( \mathcal{L} = \{ l \} \), with \( l = 1, \ldots, m \).

The association of a set of labels to each node is guided by constraints which describe the mutual consistency between labelings for n-tuples of entities in certain (spatial) arrangements. In spatial vision, locations on a discrete grid can be considered as nodes \( \mathcal{N} \) which are connected to their neighbors via the set of links \( \mathcal{E} \) (in a given connectivity structure).

**Description of the modeling framework.** In the case of boundary finding, the labeling problem can be formulated as one of finding the most likely set of orientations at the boundary positions. In order to deal with the sparsity of edge locations, the set of orientation labels is extended by a “no-line” label to account for the non-responsiveness of cells at locations that are not elements of a boundary. Initial likelihoods (or responses) are generated by oriented filters that measure the presence of local structure at different orientations in the input image. These responses need to be normalized against the maximum response in the image. This allows the treatment of filter responses as probabilities for assigning orientation labels to locations. The relatability of orientations (in the image graph) is measured by a compatibility function defined over n-tuples of orientations. In order to keep the complexity of this function at a reasonable level, only pairs of labels are evaluated. The individual strengths for orientation measures at a given image location are iteratively updated through the support that is gathered from relatable activities in a spatial neighborhood.

The elements of the relaxation model by Parent and Zucker (1989) are as follows:

1. **Spatial weighting function**
   The weighting is only specified indirectly by the requirement of having pathways of equal length along different curved shape outlines. The path lengths of neigh-
Figure 11. This figure outlines the configuration of the spatial integration field to measure compatibilities between contrast orientations measured at different locations in Parent and Zucker's (1989) model. The target cell is located at the center with matching orientation as this vertically oriented "halbert" shape. The curved sectors drawn in each hemi-field, or lobe, denote the different curvature sectors depicting the range of uncertainty for the curvature classes assigned to each tangent orientation. (This is Figure 13 from Parent and Zucker (1989), reproduced with permission.)

neighborhood support define an “extent” predicate

\[
P_{\text{length}}^{xx, \theta} = \begin{cases} 
1 & \text{if path length } ((x, \theta) \text{ to } (x', \phi)) \text{ is closer} \\
0 & \text{otherwise}
\end{cases}
\]

(compare the geometry depicted in Figure 11); here \(\theta\) again identifies the orientation of the target cell at the center, and \(\phi\) refers to the orientation of a related cell in the spatial neighborhood). This predicate function is augmented by an intrapixel length correction factor \(\text{length}(x, x')\) to compensate for any discretization effects in the orientation space. The spatial coupling is then defined by

\[
\begin{align*}
\text{weight}_{\text{left}}^{xx, \theta} &= P_{\text{length}, \text{left}}^{xx, \theta} \cdot \text{length}(x, x') \\
\text{weight}_{\text{right}}^{xx, \theta} &= P_{\text{length}, \text{right}}^{xx, \theta} \cdot \text{length}(x, x').
\end{align*}
\]

(2) Feature cooperation

The relatability of two tangent orientations is defined by a co-circularity constraint corresponding to Eqn. 7 above and is denoted here by the function \(c_{xx, \theta, \phi}\). An additional function is defined for the partitioning of the spatial neighborhood function into curved strips of shape segments (Figure 11). The rationale for this
Figure 12. This sketch outlines the configurations of spatially related, or compatible, contrast orientations. The usefulness of a curvature constraint is highlighted in the top row in which two identical contrast configurations (at points A and B) belong to the same contour segment (left) or to different segments (right). The bottom row shows a sketch of how the compatibility between curvature classes can be expressed by an intersection measure of overlap between sectors of the spatial integration fields (grey regions). Incompatible curvatures tend to reduce the area of intersection (left) while compatible ones maximize this measure (right). (These are Figures 11 and 12 from Parent and Zucker (1989), reproduced with permission).

Partition is sketched in Figure 12 (top) showing that two identical spatial configurations of tangents are related if they are part of the same boundary segment (thus having the same curvature). Conversely, if the tangents appear to belong to two segregated parts it is likely that they do not share the same curvature class \( K_k(\cdot) \). The local assignment of a curvature class is defined by an additional predicate

\[
\hat{K}_{\mathbf{x}\mathbf{x'}}^k \theta = \begin{cases} 
1 & \text{if } \rho_{\text{min}}^k \leq \rho_{\mathbf{x}\mathbf{x'}\theta}^k \leq \rho_{\text{max}}^k \text{ (radius of curvature)} \\
0 & \text{otherwise}
\end{cases},
\]  

(15)

where \( \rho_{\mathbf{x}\mathbf{x'}\theta} \) denotes an estimate of local curvature derived from local tangent orientations.

Two tangent orientations are mutually relatable if they belong to compatible curvature classes. This constraint is displayed in Figure 12 (bottom) showing the consistency of curvature classes as the overlap sections (grey regions) of the curvature sectors. This constraint is again encoded by a binary predicate function

\[
C^{kk'}_{\mathbf{x}\mathbf{x'}} \theta \hat{\theta} = \begin{cases} 
1 & \text{if } K_k(\theta) \sim K_{k'}(\hat{\theta}) \\
0 & \text{otherwise}
\end{cases},
\]  

(16)

where \( \sim \) means compatibility between curvature classes. The pairwise relatability of tangent orientations is then finally evaluated by

\[
\text{relate}_{\mathbf{x}\mathbf{x'}\theta \hat{\theta}} = c_{\mathbf{x}\mathbf{x'}\theta \hat{\theta}} \cdot \hat{K}_{\mathbf{x}\mathbf{x'}\theta \hat{\theta}}^k \cdot C^{kk'}_{\mathbf{x}\mathbf{x'}} \theta \hat{\theta}.
\]  

(17)
The input activation is generated by a field of neural line detectors that measure the presence of oriented line-like structure. Before entering the next stage of integration, their activity \( y_{x\theta} \) undergoes a stage of lateral non-maximum suppression in a discrete 3 X 3 position-orientation neighborhood, \( N_{3x3} \). This filtering can be expressed as

\[
M_{x\theta} = \begin{cases} 
1 & \text{if } y_{x\theta} \text{ is a lateral maximum} \\
0 & \text{otherwise}
\end{cases}
\]  

(18)

to form a binary label.

(3) Support and activation function

The support is computed by the combination of sub-fields and their result from integrated activities. We get

\[
z_{x\theta} = \max_{k=1,k_{max}} \sum_{x'\phi} y_{x'\phi} \cdot M_{x'\phi} \cdot \text{relate}_{xx'\theta\phi} \cdot (\text{weight}_{xx'\theta} + \text{weight}_{xx'\theta})
\]  

(19)

The final activation is computed in a stage of normalizing and remapping the support for a given target cell, i.e.,

\[
v_{x\theta} = \frac{z_{x\theta} - z_{min}}{z_{max} - z_{min}}.
\]  

(20)

This network utilizes an activation dynamics in which the normalized support values \( v_{x\theta} \) is fed back to the layer of the current input activities.\(^4\) A discrete time iterative update rule follows which can be described as

\[
\psi_{x\theta}^{i+1} = f\left( \psi_{x\theta}^{i}, v_{x\theta} \right),
\]  

(21)

where \( f(\cdot) \) denotes some function that takes into account the normalized total support for all orientations at a given location.

Variants of the model and further details. The relaxation scheme reviewed above can be considered as an elaborated version of earlier simplified schemes of relaxation labeling (e.g., Zucker, 1985; Zucker, Hummel, and Rosenfeld, 1977; Zucker and Parent, 1984). The previous models did not include any geometric definition of pairwise compatibilities between two activations, nor did they include any definition of curvature classes. The generic support function is defined as

\[
z_{x\theta} = \sum_{x'} \left( \text{weight}_{xx'\theta} + \text{weight}_{xx'\theta} \right) \cdot \sum_{x'\phi} y_{x'\phi} \cdot \text{relate}_{xx'\theta\phi}.
\]  

(22)

\(^4\)In the context of (probabilistic) relaxation labeling, the current activity of a cell selective for a given orientation denotes the probability of proper assignment of a given label (orientation) at a particular location.
Figure 13 This figure shows the processing results for a fingerprint image (left). Tangent fields are shown after two iterations of the relaxation process (right). (This is Figure 19 from Parent and Zucker (1989), reproduced with permission.)

The relatability function ‘relate’ is usually defined so that the values are in the interval [-1, 1] – with negative relatabilities for incompatible orientations and with positive relatabilities for compatible arrangements. The activation is generated by the update rule

\[ y_{x\theta}^{(t+1)} = \frac{y_{x\theta}^t [1 + z_{x\theta}]}{\sum_{\phi} y_{x\phi}^t [1 + z_{x\phi}]} \]  \hspace{1cm} (23)

In all, the activation is normalized locally against the total support gathered for different orientations. (Compare Eqns. 20 and 21; see Zucker, Hummel, and Rosenfeld, 1977). The scheme proposed by Parent and Zucker (1989) thus investigates the elaboration of the spatial relatability function by formalizing the co-circularity constraint and by incorporating the compatibility of curvature classes of pairs of orientations.

In order to make the relaxation process more sensitive to the image structure, Zucker and Parent (1984, 1985) proposed a mechanism to better evaluate the responses of oriented line detectors of different sizes and to measure the deviations of observed responses from expected ones. The expectations have been computed off-line for frequently occurring image structures. The relatability in this model is expressed as

\[ \text{relate}_{xx',\theta} = 1 - \sum_S \| y_{\theta,S}^{\text{expect}} - y_{\theta,S}^{\text{obs}} \|, \]  \hspace{1cm} (24)

where \( S \) denotes the scale, or filter size, \( \text{expect} \) represents the pre-computed model responses, \( \text{obs} \) denotes the currently measured, or observed, activities, and \( \| \cdot \| \) defines a proper norm. It is assumed that the \( y \) responses are normalized, thus having magnitudes
Figure 14. Left panel: drawing of a curved dark line on a light background. (The image is corrupted with noise). Right panel: result of the relaxation process for finding the tangents of the noisy contour after six iterations. Note that this process does not include any mechanism of lateral inhibition, or non-maximum suppression. (This is Figure 14.4 from Zucker and Parent (1985), reproduced with permission).

in the range \([0, 1]\). Identical responses between expectation and observation lead to maximum relatability, i.e. relate_{\mathbf{x}_i, \theta} = 1, whereas dissimilar responses extinguish the measure of relatability. In other schemes, e.g. in Parent and Zucker (1989) and Grossberg and Mingolla (1985), the expected visual structure is encoded in the weighting pattern of the matching field, which leads to a correlative metric to measure similarity between model expectations and observations.

**Characteristic simulations.** Applications of the relaxation labeling approach are shown in Figures 13 and 14. The first case shows the result for a noisy image of a finger print, whereas the second case shows the extraction for an isolated line on a noisy background. These examples demonstrate the reduction of uncertainty of localized measures based on the integration via an association field and also demonstrate its robustness against noise. The iterative update mechanism is designed to increase the certainty of consistent measurements against the inconsistent ones at each spatial location independent from any other support measure. This leads to a potential broadening of the response field of the grouping or curve enhancement (compare Figure 14). In order to compensate for this, the non-maximum suppression \((M_{\mathbf{x}, \theta}; \text{Eqn. 18})\) is utilized to achieve a proper thinning. Alternatively, the activity distribution can be post-processed by a separate thinning algorithm.
Filter model for the forward integration of end-stop responses

Based on the results of electrophysiological investigations of the response properties of V2 contour cells (e.g., Peterhans and von der Heydt, 1989; von der Heydt and Peterhans, 1989), Heitger et al. (1998) suggested a computational model of a mechanism for contour grouping. (See also Heitger and von der Heydt (1993) for a first description of the scheme.) Figure 15 presents an overview of the elementary processing stages. The architecture consists of a set of hierarchically organized filtering stages. Oriented contrast is measured at the stage of simple and complex cells. The result of oriented end-stop (ES) detection from complex cell responses feeds into a segregated grouping pathway. The definition of the proposed grouping mechanism is mainly influenced by the goal to make the scheme sensitive to figure-ground direction. The rationale for this design principle is the observation that partial occlusions or surface borders often generate abrupt terminations of surface items – preferentially at the side of the partially occluded surface (i.e., the background; compare also Finkel and Edelman (1989)). Also Baumann, van der Zwan, and Peterhans (1997) found evidence that cells of V2 that respond to illusory contours are sensitive to the direction of contrast that is compatible with the occlusion direction. In the computational scheme, model ES cells signal the possible directions of the occluding surface at corners and line ends. (Compare
Grossberg and Mingolla’s (1985a, 1985b) treatment of “end cuts.”) Two independent grouping rules were introduced to deal with different occlusion features, namely *ortho* grouping at line ends and *para* grouping for discontinuities in contour orientation, e.g. at corners. In order to generate a dense representations of boundary activation, complex cell responses to luminance contrasts are added to the interpolated contours between ES cells.

**Description of the modeling framework.** The elements central for the definition of the grouping model of Heitger *et al.* (1998) are the following:

1. **Spatial weighting function**
   The model utilizes the same scheme of bipolar weighting functions as those originally proposed by Grossberg and Mingolla (1985b). The weighting functions are separable in polar space such as those described in Eqn. 4 defining the spatial kernels weight\(\text{left}^{xx\theta}\) and weight\(\text{right}^{xx\theta}\).

2. **Feature cooperation**
   The *relatability* of oriented spatially separate ES responses is defined by a rule-based scheme that incorporates an implicit notation of curvature consistency. In order to distinguish between directions of figure and ground, the scheme requires that only those corners and line-ends cooperate which point in the same direction.

   *Para and ortho grouping.* Figure 16 sketches the logic of the grouping strategy. In a nutshell, para groupings occur for ES responses at corners oriented towards the location of the target grouping cell between two relatable items. The relatability is indicated by the orientation of ES cells orthogonal to the grouping direction: Same orientations cooperate, opposite orientations do not. Similarly, ortho groupings can be established between responses of like orientations, ES cells of opposite orientation are not relatable. In order to keep a specificity to figure-ground direction, separate cooperation fields are established for each individual termination direction.

   *Orientation variability and curvature.* The scheme is designed so that the association fields only integrate responses of ES cells in individual lobes of different orientations. The partitioning and combination of the integration (or association) field is a spatial variant of the curvature classes introduced by Parent and Zucker (1989). Four “curvature classes” are distinguished that are depicted in Figure 17.\(^5\) In this scheme, the integrated responses from

---

\(^5\) In the original definition a weighted sum of pairs of bipoles at different orientations is proposed (Heitger *et al.*, 1998). However, a simple rearrangement of terms leads to the scheme outlined above.
either side of a grouping cell are not simply added as in Zucker’s proposal but a non-linear AND-gate combination is used, similar to the Grossberg and Mingolla (1985b) scheme, as is next described.

The input activation to the grouping scheme is generated by a field of oriented ES cells. In order to account for any orientation variability the ES responses are summed over nearest neighbor orientations. We get

$$\tilde{w}_{\theta} = \sum_{k \in \{-1, 0, 1\}} w_{\theta+k\Delta\theta}$$

Responses assigned to ortho- and para-grouping, respectively, are computed separately for both figure-ground directions $\uparrow$ and $\downarrow$. The following equations have been proposed (for notational simplicity, we omit subscripts for location and orientation):

$$x^\uparrow_{ortho} = \tilde{w}^\uparrow, \quad x^\downarrow_{ortho} = \tilde{w}^\downarrow$$

and

$$x^\uparrow_{para} = p^\uparrow \cdot w^\downarrow, \quad x^\downarrow_{para} = p^\uparrow \cdot w^\uparrow$$

with $p^\uparrow = \tilde{w}^\uparrow/(\tilde{w}^\uparrow + \tilde{w}^\downarrow)$ and $p^\downarrow = \tilde{w}^\downarrow/(\tilde{w}^\uparrow + \tilde{w}^\downarrow) = 1 - p^\uparrow$. These responses of ortho
Figure 17. The figures show the integration (or association) field that has been partitioned into sectors of individual lobes. The association fields shown at the top support nearly straight horizontal completion allowing slight boundary curvatures to either side of the grouping cell location. The fields shown at the bottom realize grouping of left or right curvature segments. The numbers indicate the individual weights of contribution from each lobe.

and para contributions are linearly interpolated. We get the activities

\[ y^{\uparrow \text{left}} = (1 - \alpha)x^{\uparrow \text{ortho}} + \alpha x^{\uparrow \text{left}}, \quad y^{\uparrow \text{right}} = (1 - \alpha)x^{\uparrow \text{ortho}} + \alpha x^{\uparrow \text{right}} \]

where the interpolation is controlled by the “cornerness” index

\[ \alpha = 2 \sum_{\phi}(w_{\phi} \cdot w_{\phi_{\perp}})^{1/2} / \sum_{\phi}(w_{\phi} + w_{\phi_{\perp}}) \]

which gives a value \( \alpha \approx 1 \) for corners and \( \alpha \approx 0 \) for line-ends.

(3) Support and activation function

The support is computed by the multiplicative combination (AND-gating) of bipole lobes. We get

\[ z^{\uparrow}_{x\theta} = \sum_{k} \left[ \sum_{x'\phi} \left( y^{\uparrow}_{x'\phi} \cdot \text{weight}^{k\text{left}}_{xx'\theta} \right) \cdot \sum_{x'\phi} \left( y^{\uparrow}_{x'\phi} \cdot \text{weight}^{k\text{right}}_{xx'\theta} \right) \right], \quad (29) \]

were \( k \) denotes the different “curvature classes” that were represented by the differently oriented lobes (Figure 17). The grouping response for the opposite figure-ground direction \( z^{\downarrow}_{x\theta} \) are computed analoguously. The boundary representation is computed by pooling the activities of complex cells \( c \) and both grouping activations selective for opposite figure-ground directions

\[ u_{x\theta} = c_{x\theta} + m \left( z^{\uparrow}_{x\theta} + z^{\downarrow}_{x\theta} \right). \quad (30) \]

The constant \( m \) regulates the relative importance of luminance contrast and grouping strength. The responses are spatially blurred and thus positionally uncertain.
Figure 18. Contour grouping based on selective integration of ES responses using
the scheme of Heitger and coworkers. Input pattern (left) with an occlusion pat-
tern generated by a near surface and an occlusion generated by a hole in the ground
revealing a far surface (from Kanizsa, 1979). Final contour representation (right)
from oriented contrast responses and grouping of ES responses along the illusory
contours that were induced by the occlusions. Due to the selectivity of the group-
ing scheme to the orientation of ES directions, figure-ground direction is signaled
correctly (not shown here). (These panels are Figures 16 (a) and (b) from Heitger
et al. (1998), reprinted with permission.)

A stage of non-maximum suppression is applied to the pooled boundary responses
generating $u_{\theta_{\text{max}}}$. The selection of the direction of figure and ground is based on a voting scheme.
The direction of the background is assigned to the side that contributes more
ES responses in the grouping (see Figure 18). This is simply calculated by

$$D = z_{\theta_{\text{max}}} - z_{\theta_{\text{max}}}$$

and then taking the sign of the result.

Variants of the model and further details. The present grouping scheme is an elab-
aoration of an earlier one sketched by Peterhans and von der Heydt (1991), who proposed
that pairs of ES cells with distant receptive fields that signal the same end direction
are connected by gating cells. (See the description in Lesher (1995)). Gating cells of
opposite directions are pooled together with complex cells of orthogonal orientation that
signal real contrasts. The non-linear grouping of pairs of ES cell responses requires a
precise definition of the width of the connectivity pattern to span certain gap sizes. In
order to successfully group across different gap sizes, several gating cells must exist at
each spatial position that integrate ES responses from different distances. A formalized
representation of such an integration scheme (for one direction) is given by

$$z_{x}[\theta] = \sum_{o_{\text{ffset}}} y_{-o_{\text{ffset}}} \cdot y_{o_{\text{ffset}}} \cdot y_{+o_{\text{ffset}}},$$  

where $o_{\text{ffset}}$ denotes the distance from the target location along $\theta$ (the integration
for $z_{x}[\theta]$ is computed in an analogous fashion). This simplified version in Eqn. 31
Figure 19. Performance of the model for a natural scene. Input image that contains leaves which partially occlude the wires of a metal fence (left). Complete contour map of contrasts and grouping responses (right). Critical regions of contour completions at occlusions are highlighted by circles. (These panels are Figures 17 (a) and (c) from Heitger et al. (1998), reprinted with permission.)

is a special case of the scheme that utilizes a bipole mechanism (compare Eqn. 29). The summation is absorbed by the use of spatial weighting functions \( \text{weight}_{k}^{\text{left}} \) and \( \text{weight}_{k}^{\text{right}} \), respectively. Also, more flexibility has been introduced in the elaborated scheme by integration from cells of different orientations \( \phi \) and neighboring lobes of the bipole in order to deal with curved boundaries.

**Characteristic simulations.** Processing results of the filtering and grouping scheme are shown in Figs. and . The first case shows the result for two occlusion patterns in opposite figure-ground relation (left). Illusory contours are generated along the abrupt terminations that are caused by surface occlusions (middle). Figure-ground direction is correctly predicted (though not shown in Figure 19) through the voting of the directions of ES cells. The proposed voting principle that assigns the background direction according to the sign of the magnitude \( D \) (see above) is successful in many cases. However, figure-ground direction cannot always be determined by local rules alone as proposed by Heitger and colleagues. (See Lesher (1995), Figure 4b and discussion.) Figure 19 shows the application of the Heitger et al. (1998) scheme for a real world image in which partial occlusions were interpolated depending on their relative widths.

**Other models**

The models (and their variants) outlined above define different classes of approaches for feature integration based on local mechanisms. The *BCS mechanism* emphasizes the
role of feedback in order to sharpen completed contours, amplify consistent information and explain asymmetries and conflicting effects in the perception of real and illusory boundaries. The relaxation labeling approach emphasizes the enhancement of noisy measures in a consistent and salient arrangement to generate smooth curves. The process is guaranteed to converge to a stable equilibrium. The filter scheme, finally, emphasizes the fast forward integration of selected stimulus features utilizing parallel channels of selective processing and the subsequent integration of segregated representations into one final result. Below, we discuss some models that have been developed more recently by highlighting their new features and computational competences in comparison with the models discussed above.\(^6\)

\(^6\)It should be emphasized here that we do not review a number of grouping and integration models that appeared in the computational vision literature, unless they aim to deal explicitly with the explanation of perceptual data.
Completion fields

Williams and colleagues have proposed a framework that investigates the shape of grouping fields and utilized the derived weightings for simulating illusory contour generation (Thornber and Williams, 1996; Williams and Jacobs, 1997a, 1997b). The key idea is to consider the random walk of particles in a discrete lattice of the sampled space-orientation domain. The goal is to determine the probability density function for elements of a contour path between two end points (which in turn are assumed to be generated by ES neurons as in the model of Heitger et al. (1998). The stochastic process underlying the random walk is divided into the components of a source field and a sink field. Each of them denote the probability densities that a particle at a given spatial location \((x, y)\) and orientation \(\theta\) (the initial condition) reaches another point \((i, j, \phi)\) on the space-orientation lattice. The stochastic completion field is finally the product of the source and the sink field. Its strengths represent the likelihood of smooth paths connecting a pair of points (i.e., ES responses).

The strength of boundary interpolation can be computed by the spatial convolution of impulse functions on the space-orientation lattice with the kernels of the stochastic completion field. In terms of our framework developed in this chapter, this kernel denotes the weighting function of the spatial weighting and the relatability. Unlike previous formulations, here the two contributions do not occur in a separable fashion, but, instead appear as a joint distribution of likelihoods. This is due to the formulation of the underlying process as a stochastic walk on the 3-D lattice of spatial locations and orientations. The width of the spatial kernel can be controlled by the decay of particle activation in the corresponding Fokker-Planck equation. Also, the opening of the ‘figure-eight’ shape of the completion field is varied by the variance of the normal distribution which controls the “straightness” of the interpolated boundary (Williams and Jacobs, 1997a).

The combination of subfields, or lobes, is additive since the stochastic process which computes the path probabilities between two end points finally yields a symmetric convolution kernel. The support is therefore computed by the convolution of the input ES responses with the kernels derived from the stochastic completion fields. Figure 20 shows the simulation results for the generation of boundary completions for input images that contain illusory contour arrangements. The results show the final integrated edge map that is defined by complex cell responses for physical contrasts and the activities generated by the interpolation process. (Compare Heitger and von der Heydt, 1993; Heitger et al., 1998.) The generated activity distributions vary with the length...
of the boundary gaps and the shape of the interpolated contour. Note that the authors make no particular attempt to distinguish between the representation of boundaries for modal and amodal completions.

**Contour integration and segmentation**

Li (1998, 1999a, 1999b) proposed a model of spatial boundary integration in V1 based on a network of coupled oscillators that are defined by pairs of inhibitory interneurons and excitatory cells. Li's model also utilizes oriented weighting functions for lateral interaction between orientation selective cells. Each cell receives net input by laterally integrating activities in a spatially weighted neighborhood. The *combination of sub-fields*, or lobes – according to the scheme depicted in Figure 6 – is additive, leading to the *spatial weighting* weight \( \text{weight}_{xx, \theta} = \text{weight}_{xx, \theta}^{\text{left}} + \text{weight}_{xx, \theta}^{\text{right}} \). The selectivity of the excitatory integration field for *feature compatibility* measurement, relate \( \text{weight}_{xx, \theta}^{\phi} \), is similar to the scheme specified in Eqn. 6. The opening angle of the spatial weighting function is approximately ±45 deg. around the orientation axis of the target cell. An inhibitory influence is generated from cells at laterally offset positions and an orientation preference that is similar to the one of the target cell. The segregated fields of excitatory and inhibitory connectivity patterns are thus explicit representations of the feature compatibility function that has been proposed by Parent and Zucker (1989). A similar
computational model utilizing coupled oscillators was investigated earlier by Grossberg and Somers (1991). (See Ellias and Grossberg (1975) for an analysis of the dynamic properties of such networks.) This model, however, utilized only one-dimensional chains of oscillators in order to study binding properties as observed in previous physiological findings of stimulus-evoked synchronized oscillatory firing by Eckhorn et al. (1998) and Gray et al. (1989). Here, the influence of different coupling architectures on the synchronization of cell firing when probed with a single oriented bar in comparison to spatially disconnected bars was studied.

The main focus of Li’s investigations lies in the study of how V1 horizontal long-range integration could functionally account for the enhancement of co-circular contour arrangements in cluttered scenes (Li, 1998), the enhancement of texture region boundaries (Li, 1999a) and pop-out effects in visual search (Li, 1999b). Emphasis was devoted to the mathematical analysis and stability behavior of such a network of non-linear coupled oscillators. Simulation results demonstrate the functionality of, e.g., contour enhancement in artificial and natural scenes (see Figure 21). In comparison to the core models described above, Li’s model shares several properties with those of Zucker (e.g., Parent and Zucker, 1989) and the BCS (e.g., Grossberg and Mingolla, 1985b; Grossberg, Mingolla, and Ross, 1997; Ross, Grossberg, and Mingolla, 2000) and, consequently, leads to contrast contour enhancement properties as shown by previous models. The enhanced selectivity to texture boundaries, however, seems to be mainly contributed by the pronounced inhibitory action of cells with same orientation preference. These are integrated from lateral spatially offset positions but not from regions that overlap the field of mutually excitatory interactions along the axis of the target cell’s orientational preference. This provides the machinery to signal local orientation contrast, e.g., near the borders of texture regions, and also to sharpen the activity distribution along a ridge of activations.

**Boundary finding as model matching**

Neumann and coworkers have investigated the role of context effects that are generated by the arrangements of contrast responses that are generated along boundaries and that are integrated by a process of long-range interaction (Hansen et al., 2001; Neumann and Sepp, 1999). The main focus lies on the investigation of the functionality of recurrent cortico-cortical interaction. Adaptive resonance theory, or ART (Carpenter and Grossberg, 1988; Grossberg, 1980) proposes that the coincidence between a spatial distribution of activation at an early processing stage and the predictions from categorical representations of model shape patterns at a higher stage enables the weight adaptation of the connectivities between the associated areas. Similarly, inspired by
basic principles of pattern recognition, descending pathways between cortical areas are suggested to carry flexible templates which are subsequently compared with the features of the sensory input (Mumford, 1991, 1994). Following these ideas, Neumann and Sepp's approach considers initial filter responses from oriented contrast detectors as feature measurements that are fed forward to a stage of long-range integration for matching. This (model) matching, according to the description given earlier, utilizes a pattern of weighted connectivities which represent the most likely patterns of contrast orientations along boundary segments of expected shape outlines. The degree of match generates a signature whose activity is used to modulate those initial activations at the feature detection stage that are consistent with the model expectation.

In all, the model incorporates components which also appear in the three core models introduced above: The \textit{spatial weightings}, weight$^{\mathrm{left, right}}_{\mathbf{xv}/\theta}$ are similarly defined as in Eqn. 5. The \textit{feature compatibility} function incorporates an excitatory as well as an inhibitory weight field. Co-circular tangent orientations are supportive as they define a measure of feature relatability, relate$^{+}_{\mathbf{xv}/\theta}$, similar to the scheme of Parent and Zucker (1989) that was also adopted by Gove \textit{et al.} (1995). Instead of incorporating an additional symbolic predicate of curvature consistency (Parent and Zucker, 1989), the integration is supplied by a measure of non-relatability between pairs of contrast orientation, relate$^{-}_{\mathbf{xv}/\theta}$, which helps to enhance the selectivity of the support function by incorporating more global aspects of contour arrangements. The \textit{combination of sub-fields} utilizes a mechanism of disinhibition of activities generated by the individual lobes of the bipole when input is generated from \textit{both} lobes. The proposed mechanism utilizes a network of cross-channel inhibition in which equilibrium activity at the output stage
results in
\[ z_{x\theta} = LR \frac{2/\zeta + L + R}{1/\zeta + (L + R)/\zeta + LR}, \]  
with
\[ L = \sum_{x'\phi} y_{x'\phi} \cdot \text{relate}^{\text{left}}_{xx'\theta\phi} \cdot \text{weight}^{\text{left}}_{xx'\theta} \quad \text{and} \]
\[ R = \sum_{x'\phi} y_{x'\phi} \cdot \text{relate}^{\text{right}}_{xx'\theta\phi} \cdot \text{weight}^{\text{right}}_{xx'\theta}. \]

The constant \( \zeta \) controls the efficacy of the interactions between the bipole lobes. This mechanism incorporates a measure of feature conjunction using an explicit AND-gating functionality like the model proposed by Heitger et al. (1998). The support function’s functionality combines the feedforward integration (similar to Heitger et al. (1998) but without segregated feature representations) and recurrent feedback processing (similar to Grossberg and Mingolla (1985b)). This unifies different approaches that were based on different architectural principles and provides a computational theory of contour integration for surface boundary finding. Simulation results have demonstrated that the model, though it contains only a minimum number of rather simple computational stages, is able to replicate such diverse perceptual effects as, e.g., the integration of fragmented shapes, the modulation of complex cell responses by texture surround effects, and the generation of illusory contours for both Kanizsa and abutted grating illusions. Figure 22 shows the result of boundary enhancement through the action of top-down modulation by the result of contour matching.

**Long-range integration for contour detection**

Pettet et al. (1998) investigated a single layer network architecture which adaptively integrates contour paths in a field of randomly oriented distractors. Elementary items were generated by oriented Gabor patches of a given spatial frequency and constant phase (cf. Field et al. (1993)). The pop-out of an arrangement of aligned and co-oriented items was related to the increase in activation of network cells by the mutual facilitation of elements which form a (closed) contour.

The input to the stage of spatial integration is generated by a mechanism that comprises an adaptive decision-like processing element. Namely, a 0-1 activity distribution is generated where the 1’s were generated by those cells located at the position of a Gabor item with corresponding orientation. The support is subsequently evaluated by a separable function of spatial weighting and relatability. The spatial weighting is defined by a radially symmetric function of distance with Gaussian fall-off. The spatial relatability \( \text{relate}^{\text{relate}}_{xx'\theta\phi} \) is evaluated on the basis of two Gaussian functions of total curvature.
and change in curvature, respectively, of a fitting spline through a pair of cells and their oriented RFs (cf. Figure 1).

The spatial weighting, weight \( \mathbf{w} \), unlike the previously described approaches, is isotropic and thus imposes no preferential orientation for the grouping of oriented items in relation to \( \theta \). Furthermore, the model proposed by Pettet et al. (1998) assumes that lateral interaction extends over the entire field of the stimulus display. This leads to an adaptive scheme in which the efficacy of mutual interactions in the spatial integration process is adjusted to the display size.

The support is calculated simply by the sum of weighted input from cells with unit activation and the product weighting of weight \( \mathbf{w} \) and relate \( \mathbf{r} \). The activation function, finally, is defined by the non-linear dynamics of laterally coupled cells in an excitatory layer.

**DISCUSSION**

In this chapter we have introduced a framework and taxonomy to describe computational neural mechanisms of spatial integration in perceptual grouping. Based on this conception we presented key models as well as some derived from them using this taxonomy. It turns out that many similarities exist between several approaches, but also some differences that need to be further evaluated in order to quantify their functional significance.

All models may be coarsely classified according to the architectural principles involved in the generation of spatial integration and grouping, namely feedforward and feedback mechanisms. A related question refers to the mechanisms that help to establish the neural wiring underlying the observed functionality of grouping. The following questions turn out to be of central importance in order to reveal the principles and mechanisms of spatial long-range integration mechanisms in early and mid-level vision:

1. What are the core principles underlying the establishing of spatial integration and grouping?

2. What are the constraints and underlying mechanisms for establishing long-range interactions?

**Core principles of grouping**

Concerning the first issue, it can be stated that there still is an ongoing debate on the role of feedforward and feedback mechanisms involved in spatial grouping and
the dominance of their individual contributions in visual processing. Advocates of a feedforward-dominated principle often argue on the basis of time constraints of the brain mechanisms involved compared with the observed physiological data. For example, the time course of generating illusory contour responses related to stimulus onset argues in favor of a mechanism that is mainly feed-forward by simply integrating segregated local items, according to von der Heydt and Peterhans (1989) and Peterhans and von der Heydt (1989). This view has been further fueled by more recent experimental investigations of Thorpe et al. (1996) who investigated the lower bounds on time needed by subjects to signal the presence of an object category in a visual scene.

On the other hand, proponents of the view that feedback mechanisms play a major role in grouping begin by admitting that relatively noise-free tasks can support “one-pass,” essentially feedforward processing within an architecture capable of feedback, and then go on to argue on that recurrent processing allows more flexibility and robustness in shape boundary finding that relies on noisy estimates (Grossberg et al. 1997; Zucker, 1985). For example, Grossberg (1994) has argued that feedback mechanisms help to disambiguate and further tune the activities generated to outline the surface of a fragmented shape. Recently, Neumann and Sepp (1999) have combined these seemingly diverging architectural frameworks and have shown that grouping can occur in a “rough-and-ready” fashion by a feedforward mechanism. If there are several initial candidate groupings, feedback helps to further tune and sharpen the spatial responses based on more global information and stimulus context.

The investigation of the differences in the temporal persistence of illusory and contrast boundaries further argues in favor of feedback mechanisms (Francis et al., 1994; Francis and Grossberg, 1996). It was demonstrated empirically that the occluding surface region of a Kanizsa square (which is delineated by illusory contours) has a longer perceptual persistence in comparison to a region that is bounded by a physical contrast (Meyer and Ming, 1988). The neural representation of surface regions seems to be very similar irrespective of the type of boundary Mendola et al. (1999). It is therefore reasonable to assume that the main effect of perceptual persistence is due to the mechanisms of boundary generation. It was suggested that boundary activation that is fed by contrast detectors is immediately switched off after withdrawal of the contrast input, whereas the illusory boundaries still receive recurrent facilitation by top-down activation and therefore cause the boundary activation to persist longer Francis et al. (1994). In further support of the feedback hypothesis, Ringach and Shapley (1996) found evidence for two masking regimes for stimuli containing illusory contours, the second of which extended more than 300 msec beyond stimulus onset.
Aspects of temporal coding principles based on oscillator mechanisms or spiking neurons may play another important role in grouping tasks. Several empirical studies (Eckhorn et al., 1988; Engel et al., 1991; Gray et al., 1989) indicate that distributed representations of related scene fragments are linked by temporally correlated, or synchronized, neural activation. It is claimed that perceptual grouping based on temporal binding implements the principles of Gestalt grouping allowing the desired flexibility in the dynamic linking and segregation that alternative mechanisms such as convergent coding principles or population coding seem to lack (Gray, 1999). Several studies demonstrate that γ-band activity (30-60 Hz) signals the relatedness of stimulus features in a robust fashion (Frien et al., 1994), although this property holds only over small cortical distances of 3-5 millimeters (Saam and Eckhorn, 2000). The temporal coding hypothesis studied in isolation appears, however, to be incomplete. The temporal establishment of grouping addresses the signaling of binding, but not the "how" or "what" of its computation (Shadlen and Movshon, 1999). These questions form the core of this chapter. In all, therefore, our contribution provides a necessary framework to classify several approaches and also help to identify potentially new mechanisms to further improve the computational as well as the explanatory power of grouping models.

Establishing long-range interactions

Early in this chapter we introduced the underlying anatomical, physiological and behavioral basis of grouping. On the behavioral level, it has been demonstrated that the capability of organizing visual patterns, as in the case of illusory contours, is developed postnatal within the first seven months (Bertenthal et al., 1980).

Oriented long-range connections seem to provide the basis for smooth boundary grouping. The oriented fiber connections of an oriented target cell link with other cells of similar orientation preference that are arranged along a line defined by the target cell's orientation (see Figure 4). It remains, however, as an open issue, how these long-range interactions are established and how the stimulus properties of real world scenes constrain the parameters of the relatability function. Possible constraints and underlying mechanisms for establishing oriented lateral long-range connections have been modeled on different levels:

1. Principles of stochastic motion are utilized to investigate the oriented diffusion of particles by Williams and Jacobs (1997a). The key to this modeling approach...

---

7This observation seems to run against the results of more recent studies where babies were shown to group items into units according to the Gestalt laws of organization (Farroni et al., 2000). The stimuli used in both studies differed significantly, however: Whereas the first investigation utilized Kanizsa figures where illusory boundaries need to be generated in order to form surface-related units, the latter study used square items of different luminances that can be spontaneously grouped into columns.
lies in the computation of likely paths that interpolate between pairs of separated items. This basic idea can be generalized, however, such that a neural growth process tends to send axonal arbors along a direction tangential to the target cell’s orientation. The growth may be driven along a gradient that decreases with distance and deviation from the reference orientation.

(2) The statistical nature of oriented contrast and boundaries in natural scenes has been investigated to derive coefficients for the co-occurrence of separated pairs of oriented items (Geisler et al., 2001). These authors evaluated the statistics of the triplets defined by the fundamental quantities that were displayed in Figure 1 (bottom). The results demonstrate that the most frequently occurring contrast orientations (at given distances and orientation differences from a target cell) resemble a structure similar to the dipole icon. This information provides a key to understanding the underlying structure and extent of long-range connection utilized for grouping and integration that is based on the statistics of natural scenes. The importance of investigating the co-occurrence of oriented structure in natural scenes is further highlighted by related studies. For example, Krueger (1998) studied the second order statistics of responses generated by oriented filters that resemble the structure of cortical simple cells. The results demonstrate a significantly increased likelihood for the joint occurrence of collinear as well as parallel structure in images of real scenes. This concurs with the structure of the grouping field derived by Geisler et al. (2001). More recently and in a similar vein, Sigman et al. (2001) investigated the probability of co-occurrence of responses from oriented filters along a co-circular path. A correlation energy measure again shows that this is indeed the case, adding further evidence for the significance of long-range correlations between oriented contrast items in natural scenes.  

(3) The mechanisms of neural self-organization have been modeled in order to demonstrate that the long-range connections can be established in a stable fashion (Grossberg and Williamson, 2001). Their results directly relate to the observation discussed before that fragments of boundary structures appear with certain probabilities in natural scenes. The model investigates the stable development of neural layer 2/3 connections in cortical areas V1 and V2.

---

8It is worth adding a note about some simplifying assumptions that have been utilized in order to derive the probability measures. All approaches mentioned above utilize some sort of a binarization process in order to reduce the computational load for getting the resulting statistics. These processes, however, can in principle vary the structure of the probability distributions. It might, therefore, be useful to utilize the full power spectra instead and measure co-occurrences on the basis of oriented energy (compare Keil and Cristobal (2000)).
Figure 23. The left panel represents the earliest graphical representation of a bipole icon, or grouping support function, in the literature. Here the lengths of line segments graphically code the coupling strength between an array of filters centered at the positions of particular segments and having the orientational preferences indicated by those segments, and the bipole cell, which has a horizontal orientational preference and a receptive field center at the center of the diagram. The panel on the right uses a grayscale code, where darker lines stand for higher coupling strengths, and a similar position and orientation code. The represented values were derived by a computational and psychophysical tour de force. Although the precise values vary, the overall distributions of the two figures, one derived intuitively years ago and the other created more recently by careful measurement, are in remarkable agreement with respect to the dimensions described in Figure 1. (The left panel is Figure 32a of Grossberg and Mingolla (1985b) and the right is Figure 3e of Geisler et al. (2001); both are reproduced with permission.)

In all, these different lines of investigation shed light on the nature of the processing problems in grouping and long-range integration as well as the adaptive mechanisms to adjust the neural machinery to solve the problem of binding together separate visual elements.

CONCLUSION

The core of this chapter focuses on the computational mechanisms of spatial long-range integration. In our discussion in the previous section we identified the topics of key investigation, namely, the computational role of different principles in neural architecture, the precise coding of spatial relatedness or binding, as well as the self-organization of the wiring patterns. We have demonstrated that the common framework of all computational models is based on some variation of the structure of the bipole kernel (depicted in Figure 1) and a number of additional computational principles in evaluating related activities in feature space. The bipole kernel itself graphically visualizes the oriented fan-like connectivity between sites in a space-orientation feature space. Its initial definition was based largely on the investigation of geometrical concepts, namely the differential geometry of plane curves. The particular shape and its parametrization, however, was
derived largely by intuition. Figure 23 (left) reviews the structure of the fan-in efficacies in the original bipole kernel proposed by Grossberg and Mingolla (1985b). It shows the three-dimensional weighting kernel in the space-orientation domain for a horizontally oriented target cell.

Later, several investigations shed some light on the underlying structure of such long-range mechanisms, both structurally and functionally. Bosking et al. (1997), for example, demonstrated the selective connectivity of axon fiber connections in striate cortex. Functionally, Field et al. (1993) derived the bipole distribution of a grouping support function on the basis of their psychophysical investigation of the conditions for pop-out of spatial arrangements of oriented items in a field of distractors. Only recently, these different aspects in model definition and related empirical methods received a grounding through the investigation of the statistics of visual scene properties. Geisler et al. (2001) measured the co-occurrence of oriented contrasts in natural scenes. It turns out that the frequencies of the naturally occurring pairs of edges (at given distances and orientations) resemble the structure of the bipole kernel (compare Figure 23, right).

The findings of Geisler et al. (2001) "justify" the definition of such a kernel structure based on the statistics of the natural environment. They further provide constraints for proper parameterizations of bipole mechanisms. The plain statistics of joint co-occurrence of contrast edges in static images may only be a first step in revealing the underlying basis of grouping. Since objects and surfaces in the natural environment occur in a depth order, partially occluding each other and being in motion, the next step may be to investigate statistical distributions of even higher order, including feature dimensions such as disparity and motion direction. In all, this demonstrates that there is still a need for ongoing investigation in order to explain the principles and mechanisms of binding contour fragments and to explain how these "units" will be linked into surfaces and objects involving later stages of processing.

ACKNOWLEDGMENTS

The authors thank Thorsten Hansen, Matthias Keil, and Gurumurthy Swaminathan for helpful comments on earlier drafts, and Cynthia Bradford for assistance in manuscript preparation. H.N. was supported in part by a grant of the Volkswagen-Stiftung (VW I/75967) and E.M. was supported in part by the Defense Advanced Research Projects Agency and the Office of Naval Research (ONR N00014-95-1-0409). Email: hneumann@neuro.informatik.uni-ulm.de, ennio@cns.bu.edu
REFERENCES


Heitger, F. and von der Heydt, R. (1993). A computational model of neural contour processing: Figure-ground segregation and illusory contours. Proc. 4th Int. Conf. on Computer Vision (ECCV '93), Berlin, Germany, 32-40.


From Fragments to Objects


