LETTER **Communicated by Steven Zucker**

Neural Mechanisms for the Robust Representation of Junctions

Thorsten Hansen
thorsten.hansen@psychol.uni-giessen.de
Giessen University, Department of Psychology, D-35394 Giessen, Germany

Heiko Neumann
hneumann@neuro.informatik.uni-ulm.de
Ulm University, Department of Neural Information Processing, D-89069 Ulm, Germany

Junctions provide important cues in various perceptual tasks, such as the determination of occlusion relationships for figure-ground separation, transparency perception, and object recognition, among others. In computer vision, junctions are used in a number of tasks, like point matching for image tracking or correspondence analysis. We propose a biologically motivated approach to junction representation in which junctions are implicitly characterized by high activity for multiple orientations within a cortical hypercolumn. A local measure of circular variance is suggested to extract junction points from this distributed representation. Initial orientation measurements are often fragmented and noisy. A coherent contour representation can be generated by a model of V1 utilizing mechanisms of collinear long-range integration and recurrent interaction. In the model, local oriented contrast estimates that are consistent within a more global context are enhanced while inconsistent activities are suppressed. In a series of computational experiments, we compare junction detection based on the new recurrent model with a feedforward model of complex cells. We show that localization accuracy and positive correctness in the detection of generic junction configurations such as L- and T-junctions is improved by the recurrent long-range interaction. Further, receiver operating characteristics analysis is used to evaluate the detection performance on both synthetic and camera images, showing the superior performance of the new approach. Overall, we propose that nonlocal interactions implemented by known mechanisms within V1 play an important role in detecting higher-order features such as corners and junctions.

1 Introduction and Motivation

Corners and junctions are points in the image where two or more edges join or intersect. Whereas edges lead to variations of the image intensity along a

single direction, corners and junctions are characterized by variations in at least two directions. In other words, edges are intrinsically one-dimensional signals, whereas corners and junctions are intrinsically two-dimensional signals. Compared to regions of homogeneous intensity, edges are rare events. Likewise, compared to edges, corners and junctions are rare events of high information content. Moreover, corners and junctions are invariant under different viewing angles and viewing distances. Both the sparseness of the signal and the invariance under affine transformations and scale variations establish corners and junctions as important image features. Points of intrinsically two-dimensional signal variations such as corners and junctions have also been termed *keypoints* (Heitger, Rosenthaler, von der Heydt, Peterhans, & Kübler, 1992; Michaelis, 1997) or *interest points* (Schmid, Mohr, & Bauckhage, 2000).

Corners and junctions are useful for various higher-level vision tasks such as the determination of occlusion relationships, matching of stereo images, object recognition, and scene analysis. The importance of corner and junction points for human object recognition has been demonstrated in a number of psychophysical experiments (Attneave, 1954; Biederman, 1985, 1987). Biederman showed that object perception of line drawings is severely impaired when corners (i.e., contours of high curvature) are removed, but largely preserved when contours of low curvature are deleted. Junctions also seem to play an important role in the perception of brightness and transparency (Adelson, 1993, 2000; Metelli, 1974; Todorović, 1997). Recently Rubin (2001) proposed that local occlusion cues as signaled by junctions are necessary to trigger modal and amodal surface completion. Rubin showed that other cues, such as surface relatability and surface similarity, did not lead to the perception of illusory contours or amodal completion when junction cues are removed from otherwise unchanged stimuli.

In physiological studies in monkey visual cortex cells have been reported that selectively respond to corners and line ends (Hubel & Wiesel, 1968). Das and Gilbert (1999) showed that correlated activities of V1 cells can signal the presence of smooth outline patterns as well as patterns of orientation discontinuity as occurring at corners and junctions. Cells preferentially responding to curves and angles have been found in area V4 (Pasupathy & Connor, 1999, 2001).

Recently McDermott (2001, 2002) studied the performance of human observers for the detection of junctions in natural images. He found that the ability to detect junctions is severely impaired if subjects view the location of a possible junction through a small aperture. Detection performance and observers' confidence ratings decreased with decreasing size of the aperture. The results suggest that a substantial number of junctions in natural images cannot be detected by local mechanisms.

In this article, we propose a new mechanism for corner and junction detection based on a distributed representation of contour responses within a model hypercolumn (Zucker, Dobbins, & Iverson, 1989). Unlike local ap-

proaches as proposed in computer vision (e.g., Harris, 1987; Mokhtarian & Suomela, 1998; Parida & Geiger, 1998), the new scheme is based on a more global, recurrent long-range interaction for the coherent computation of contour responses. Such nonlocal interactions evaluate local responses within a more global context and generate a robust contour representation. A measure of circular variance is used to extract corner and junction points at positions of large responses for more than one orientation.

The letter is organized as follows. In section 2, we present the model of recurrent collinear long-range interactions and detail the new junction detection scheme. Simulation results for a number of synthetic and real-world camera images are presented in section 3. A discussion of the results is given in section 4. Section 5 concludes the article. A short version reporting this research has been published in Hansen and Neumann (2002).

2 A Neural Model for Corner and Junction Detection

Corner and junction configurations can be characterized by significantly increased responses for two or more orientations at a particular location in the visual space. A cortical hypercolumn is the neural representation for oriented responses at a particular location. Corners and junctions are thus characterized by significant activity of multiple neurons within a hypercolumn, as proposed by Zucker et al. (1989).

Multiple oriented activities as measured by a simple feedforward mechanism are sensitive to noisy signal variations. In previous work, we have proposed a model of recurrent collinear long-range interaction in the primary visual cortex for contour enhancement (Hansen & Neumann, 1999, 2001). During the recurrent long-range interactions, the initially noisy activities are evaluated within a larger context. In this recurrent process, only coherent orientation responses are preserved—responses supported by responses in the spatial neighborhood—while other responses are suppressed. Besides the enhancement of coherent contours, the proposed model also preserves multiple activity at corners and junctions. Corners and junctions are thus implicitly characterized by a distributed representation of high multiple activity within a hypercolumn.

Such a distributed representation may suffice for subsequent neural computations. However, at least for the purpose of visualization and comparison to other junction detection schemes, an explicit representation is required. Following the above considerations, corners and junctions can be marked if multiple orientations are active and high overall activity exists within a hypercolumn.

In the following, we first present the proposed model of collinear recurrent long-range interactions in V1 and then detail a mechanism to explicitly mark corner and junction points. The analysis and experiments are restricted to patterns consisting of straight or extremely low curvature boundaries.

2.1 Coherent Contour Representation by a Model of Collinear Recurrent Long-Range Interaction in V1. The model of collinear long-range interactions in V1 is motivated by biological mechanisms. The core mechanisms of the model include localized receptive fields (RFs) for oriented contrast processing, interlaminar feedforward and feedback processing, cooperative horizontal long-range integration, and lateral competitive interactions.

The key properties of the model are motivated by empirical findings:

- *Horizontal long-range connections*. The grouping of aligned contours requires a mechanism that links cells of proper orientation over larger distances. Horizontal long-range connections found in the superficial layers of V1 may provide such a mechanism. They span large distances (Gilbert & Wiesel, 1983; Rockland & Lund, 1983) and selectively link cells with similar feature preference (Gilbert & Wiesel, 1989) and collinear aligned RFs (Bosking, Zhang, Schofield, & Fitzpatrick,1997; Schmidt, Goebel, Löwel, & Singer, 1997). Evidence for nonlocal integration also comes from psychophysical experiments for contrast detection (Kapadia, Ito, Gilbert, & Westheimer, 1995; Kapadia, Westheimer, & Gilbert, 2000; Polat & Sagi, 1993, 1994) and contour integration (Field, Hayes, & Hess, 1993; Yen & Finkel, 1998). Since it is known that facilitation exists that is strong for collinear flankers, here we focus on interactions between collinear, iso-oriented items. The more general case of mutual facilitation for cocircular arrangements is considered in Parent and Zucker (1989) and Zucker et al. (1989).
- *Short-range connections*. Short-range connections are rather unspecific for a particular orientation (Amir, Harel, & Malach, 1993; Bosking et al., 1997; DeAngelis, Freeman, & Ohzawa, 1994) and most likely belong to an inhibitory system (Kisvarday, Kim, Eysel, & Bonhoeffer, 1994).
- *Modulating feedback*. Several physiological studies indicate that feedback projections have a modulating or gating rather than generating effect on cell activities (Hirsch & Gilbert, 1991; Hupé et al., 1998; Salin & Bullier, 1995). Feedback alone is not sufficient to drive cell responses (Sandell & Schiller, 1982), and initial bottom-up activity is necessary to generate activity.

The model architecture is defined by a sequence of preprocessing stages and a recurrent loop of long-range interaction, realizing a simplified architecture of V1 (see Figure 1).

Processing within the recurrent loop defines a functional architecture of two interacting regions, each with a distinctive purpose (Hansen, Sepp, & Neumann, 2001; Neumann & Sepp, 1999): The lower region serves as a stage of feature measurement and signal detection. The higher region represents expectations about visual structural entities and context information to be matched against the incoming data carried by the feedforward pathway.

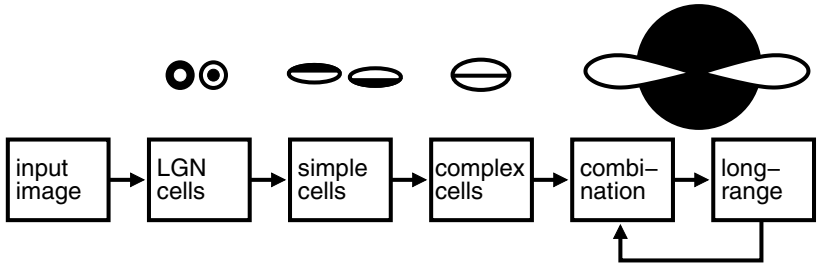

Figure 1: Overview of model stages together with a sketch of the sample receptive fields of cells at each stage for 0 degree orientation. For the long-range stage, the spatial weighting function of the long-range filter is shown together with the spatial extent of the inhibitory short-range interactions (sketched as black disk).

Feedback from the higher region then selectively enhances those activities of the lower region that are consistent with the top-level expectations or a priori assumptions (Carpenter & Grossberg, 1988; Mumford, 1991).

2.1.1 Feedforward Preprocessing. In the feedforward path, the initial luminance distribution I is processed by isotropic lateral geniculate nucleus (LGN) cells $K_{\text{on/off}}$, followed by orientation-selective simple cells S and complex cells C (see Figure 1). The interactions in the feedforward path are governed by basic linear equations to keep the processing in the feedforward path relatively simple and to focus on the contribution of the recurrent interaction. In our model, complex cell responses C provide an initial local estimate of contour strength, position, and orientation, which is used as bottom-up input for the recurrent loop. The equations that govern the computations in the feedforward path are detailed in appendix A.

2.1.2 Recurrent Long-Range Interaction. The output of the feedforward preprocessing defines the input to the recurrent loop. The recurrent loop has two stages: a combination stage, where bottom-up and top-down inputs are fused, and a stage of long-range interaction.

The equations that govern the interaction at both the combination stage and the long-range stage employ nonlinear, multiplicative interactions. The divisive inhibition resulting from this interaction is sometimes referred to as shunting inhibition (Furman, 1965; Grossberg, 1970; Hodgkin, 1964; Levine, 2000). Shunting inhibition has been proposed as a possible mechanism for divisive interactions in a number of studies (Borg-Graham, Monier, & Frégnac, 1998; Carandini & Heeger, 1994). However, a number of results, in particular from modeling studies, have challenged this view, proposing that shunting inhibition has a subtractive rather than divisive effect on cell firing

rates (Douglas & Martin, 1990; Holt & Koch, 1997). More recent studies have shown that shunting inhibition can indeed divisibly modulate firing rates in neurons with high-variability synaptic inputs and dendritic saturation (Mitchell & Silver, 2003; Prescott & Koninck, 2003). Based on these results, we propose that shunting inhibition is a physiological plausible candidate mechanism for the divisive inhibition as used in the model.

Combination stage. At the combination stage, feedforward complex cell responses C and feedback long-range responses W are added and subject to a nonlinear compression of high-amplitude activity following the Weber-Fechner law (Fechner, 1889; Weber, 1905):

$$V = \beta_V \frac{\text{net}_\theta}{\alpha_V + \text{net}_\theta}, \qquad \text{where } \text{net}_\theta = C + \delta_V W. \tag{2.1}$$

This equation is the steady-state response $\partial_t V = 0$ of the differential equation

$$\partial_t V = -\alpha_V V + (\beta_V - V)\ \text{net}_\theta. \tag{2.2}$$

The divisive inhibition resulting from this interaction yields a bounded activity in the range $[0, \beta_V]$. More precisely, the activity V is described by

$$\begin{aligned} V = V(\text{net}_\theta) &= \beta_V \frac{\text{net}_\theta}{\alpha_V + \text{net}_\theta} \\ &= \begin{cases} 0 & \lim_{\text{net}_\theta \to 0} \\ \frac{n}{1+n}\,\beta_V & \text{net}_\theta = n\alpha_V, \quad n \in \mathbb{R}^+ \\ \beta_V & \lim_{\text{net}_\theta \to \infty}. \end{cases} \end{aligned} \tag{2.3}$$

Plots of the function defining V for different values of the decay parameter α_V and a unit-valued scaling parameter $\beta_V = 1$ are depicted in Figure 2.

The weighting parameter $\delta_V = 2$ in equation 2.1 is chosen so that dimensions of C and W are approximately equal, the decay parameter $\alpha_V = 0.2$ is chosen small compared to net_θ, and the scaling parameter $\beta_V = 10$ scales up the activity to be sufficiently large for the subsequent long-range interaction. In the first iteration step, feedback responses W are set to C.

Long-range interaction. At the long-range stage, the contextual influences on cell responses are modeled. Orientation-specific, anisotropic long-range connections provide the excitatory input. The inhibitory input is given by isotropic interactions in both the spatial and orientational domains. Long-range connections are modeled by a filter whose spatial layout is similar to the bipole filter as first proposed by Grossberg and Mingolla (1985a). The spatial weighting function of the long-range filter is narrowly tuned

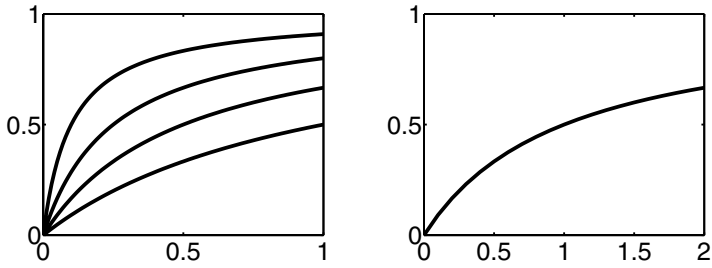

Figure 2: Compressive nonlinearity of the function $y = x/(x + \alpha)$ employed in equation 2.1. (Left) Plot of the compression function for various values of the decay parameter $\alpha = 0.1, 0.25, 0.5$, and 1 (Right) The decay parameter α in equation 2.1 can be viewed as defining the unit length of the input values x in the compression function; input values x given as a multiple of α, that is, $x = n\alpha$, are mapped to $n/(1 + n)$ independent of α. Scaling of the x-axis by $1/\alpha$ thus maps all y-values as obtained for different values of α on a single curve given by $y = n/(1 + n)$.

to the preferred orientation, reflecting the highly significant anisotropies of long-range fibers in visual cortex (Bosking et al., 1997; Schmidt et al., 1997). The size of the long-range filter is about four times the size of the RF of a complex cell as sketched in Figure 1.

Essentially, excitatory input is provided by correlation of the feedforward input with the long-range filter B_θ. A cross-orientation inhibition prevents the integration of cell responses at positions where responses for the orthogonal orientation also exist. The excitatory input is governed by

$$L_\theta = [V_\theta - V_{\theta^\perp}]^+ \star B_\theta, \tag{2.4}$$

where $\star$ denotes spatial correlation and $[x]^+ = \max\{x, 0\}$ denotes half-wave rectification.

The long-range filter is defined as a polar-separable function:

$$B_\theta(\varphi, r) = B_{\text{ang}}(\varphi) B_{\text{rad}}(r). \tag{2.5}$$

The angular function B_{ang} is maximal for the preferred direction θ and smoothly rolls off in a cosine fashion, being zero for angles deviating more than $\alpha/2$ from the preferred orientation:

$$B_{\text{ang}}(\varphi) = \begin{cases} \cos(2\pi \frac{1}{2\alpha}(\theta - \varphi)) & |\varphi - \theta| \le \alpha/2 \\ 0 & |\varphi - \theta| > \alpha/2. \end{cases} \tag{2.6}$$

The parameter α, which defines the opening angle of the long-range filter, is set to 20 degrees. The radial function B_{rad} is constant for values smaller

Figure 3: Spatial weighting function for the long-range interaction for a reference orientation of 0 degrees.

than $r_{\max} = 25$ and smoothly decays to zero in a gaussian fashion for values larger than $r_{\max}$:

$$B_{\text{rad}}(r) = \begin{cases} 1 & r \leq r_{\max} \\ \exp(\frac{-r^2}{2\sigma}) & r > r_{\max}. \end{cases} \tag{2.7}$$

The standard deviation of the gaussian is set to $\sigma = 3$. The long-range filter is finally normalized such that the filter integrates to one. A plot of the long-range filter for a reference orientation of 0 degrees is depicted in Figure 3.

Responses are not salient if nearby cells of similar orientation preference also show a strong response. Therefore, an inhibitory term is introduced that gathers activity from the spatio-orientational neighborhood of the target cell. Gaussian weighting functions are used to sample activity in both space (x, y) and orientation domain θ. The inhibitory input M to the long-range interaction is thus provided by the three-dimensional (3D) correlation of the input L with a 3D gaussian $G_{\boldsymbol{\sigma}}$:

$$M(\mathbf{x}) = G_{\boldsymbol{\sigma}}(\mathbf{x}) \ \star \ L(\mathbf{x}). \tag{2.8}$$

Since the 3D gaussian $G_{\boldsymbol{\sigma}}(\mathbf{x})$ is separable in each dimension, the 3D correlation can be efficiently implemented as the successive correlation with three one-dimensional (1D) gaussians $g_{\sigma_{\text{sur}}}(x)$, $g_{\sigma_{\text{sur}}}(y)$ and $g_{\sigma_\theta}(\theta)$. Since correlation is associative, the correlation can also be conceptualized as two successive correlations: a two-dimensional (2D) correlation in the spatial domain with an isotropic 2D gaussian $G_{\sigma_{\text{sur}}}$ followed by a correlation in the orientational domain with a 1D gaussian g_{σ_o}:

$$M(\mathbf{x}) \equiv M(x, y, \theta) \ = \ L(x, y, \theta) \ \star \ G_{\sigma_{\text{sur}}}(x, y) \ \star \ g_{\sigma_o}(\theta). \tag{2.9}$$

The standard deviation of the gaussian in the spatial domain is set to $\sigma_{\text{sur}} = 8$ to model the smaller extent of the inhibitory short-range connections. This parameterization results in an effective spatial extension of about half the size of the excitatory long-range interaction modeled by the long-range filter. The standard deviation in the orientational domain is set to $\sigma_o = 0.5$ to give near-zero input for the orthogonal orientation. The 3D gaussian kernel is visualized in Figure 4.

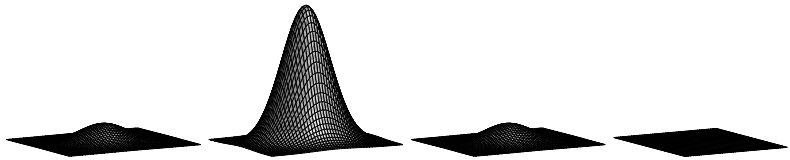

Figure 4: 3D gaussian in (x, y, θ)-space visualized as $O_{\max} = 4$ gaussians in (x, y) space. (Left to right) Gaussians for $\theta - \Delta\theta$, θ, $\theta + \Delta\theta$, and $\theta + 2\Delta\theta$. Note that the gaussian for the orientation $\theta_\perp = \theta + 2\Delta\theta$ orthogonal to θ is near zero at all positions.

The final activity of the long-range stage results from interactions between the excitatory long-range input L and the inhibitory input M. The excitatory long-range input L is gated by the activity V to implement a modulating rather than generating effect of lateral interaction on cell activities (Hirsch & Gilbert, 1991; Hupé et al., 1998). Similar to equation 2.1, the inhibition from M is divisive. The final activity of the long-range stage thus reads

$$W = \beta_W \frac{V(1 + \eta^+ L)}{\alpha_W + \eta^- M}, \tag{2.10}$$

where $\alpha_W = 0.2$ is the decay parameter and $\eta^+ = 5$, $\eta^- = 2$, and $\beta_W = 0.001$ are scale factors. This equation is the steady-state response $\partial_t W = 0$ of the differential equation

$$\partial_t W = -\alpha_W W + \beta_W V(1 + \eta^+ L) - \eta^- W M.$$

The result of the long-range stage is fed back and combined with the feedforward complex cell responses in equation 2.1, thus closing the recurrent loop. The multiplicative interactions governing both the long-range interactions and the combination of feedback and feedforward input ensure rapid equilibration of the dynamics after a few recurrent cycles and result in graded responses within a bounded range of activations.

The model is robust against parameter changes mainly caused by the compressive transformation equations employed. For the combination of responses (see equation 2.1), however, it is crucial to have activities in both streams of similar order of magnitude. Also, the relative RF sizes must not be substantially altered. The current parameter setting results in relative RF sizes of complex cells: isotropic inhibitory short-range filter: long-range interaction of about 1:2.5:4, assuming a cut-off of the gaussians at 2σ (or 95% of the total energy).

2.2 Junction Detection by Readout of Distributed Information Using a Measure of Circular Variance. As stated above, corners and junctions are characterized by points in the visual space where responses for multiple orientations are present and high overall activity exists within a hypercolumn. For the readout of this distributed information, a measure of circular variance is used to signal multiple orientations. The overall activity is given by the sum across all orientations within a hypercolumn. Thus, the junction map J for a distributed hypercolumnar representation such as the activity of the long-range stage W (see equation 2.10) is given by

$$J = \text{circvar}(\mathrm{W}_\theta)^2 \sum_\theta \mathrm{W}_\theta. \tag{2.11}$$

The junction map J thus receives high values from a combined measure of high activations in multiple orientation channels and high overall activity. The function "circvar" is a measure of the circular variance within a hypercolumn. The squaring operation enhances the ratio between high and low circular variances. Circular variance is defined as one minus the normalized population vector

$$\text{circvar}(\mathrm{W}_\theta) = 1 - \frac{|\sum_\theta \mathrm{W}_\theta \exp(2\mathrm{i}\theta)|}{\sum_\theta \mathrm{W}_\theta}. \tag{2.12}$$

Circular variance takes values in the range [0; 1]. A value of 0 denotes a single response, whereas a value of 1 occurs if all orientations have the same activity. Circular variance has been used in a number of physiological studies to characterize the response properties of cells in V1 (McLaughlin, Shapley, Shelley, & Wielaard, 2000; Pugh, Ringach, Shapley, & Shelley, 2000; Ringach, Hawken, & Shapley, 1997).

To visualize the data, single junction points are marked as local maxima in the junction map. First, the junction map is smoothed with a gaussian ($\sigma = 3$) to regularize the junction map and improve the localization accuracy. Corner points are then marked as local maxima whose strength must exceed a fraction of the maximum response in the smoothed junction map. Local maxima are computed within a 3×3 neighborhood. These operations help to visualize the junction data and are not meant to have a direct physiological counterpart.

The corner detection scheme detailed in this section reads out a scalar value of junction activity from the population vector of orientation responses within a hypercolumn. The readout operation does not rely on any specific properties of the long-range stage and thus can be applied to any input where multiple orientations are locally represented within a hypercolumn. To demonstrate the advantages of the recurrent long-range interaction as opposed to a purely feedforward processing, it is instructive to compare the detection results obtained for two different kinds of distributed input:

the long-range activity W and the complex cell activity C. Results of this comparison are presented in the next section.

3 Simulations

In this section, we show the competencies of the proposed junction detection scheme for a variety of synthetic and real-world camera images. In particular, the robustness to noise and the localization properties of the new scheme are evaluated. In order to focus on the relative merits of the recurrent long-range interactions for the task of corner and junction detection, the proposed scheme is evaluated using two different kinds of input: the activity W of the long-range stage and the purely feedforward activity C of the complex cell stage. Model parameters as specified in section 2 are used in all simulations.

The multiplicative interactions that govern the computations at the long-range stage and the combination stage ensure an equilibration of the model dynamics after a small number of recurrent cycles. For all input images, model dynamics have saturated after 12 recurrent cycles. The model responses continuously improve with each iteration, with the largest improvement achieved in the first interactions (Hansen, 2003). A fixed number of recurrent cycles is therefore not crucial for the model performance. This is of particular relevance for real biological networks, where the number of recurrent interactions is limited by the time available to process a stimulus.

3.1 Localization of Generic Junction Configurations. From the outset of corner and junction detection in computer vision, the junction types have been partitioned into distinct classes like T-, L-, and W-junctions (Huffman, 1971). This catalog of junction configurations has been extended by other junction types such as Ψ-junctions, which have been suggested to provide strong cues for inferring surface shading and reflectance (Adelson, 2000; Sinha & Adelson, 1993).

In the first simulation, we compare the localization accuracy of junction responses based on feedforward versus recurrent long-range responses for L-, T-, Y-, W- and Ψ-junctions (see Figure 5). To evaluate the localization accuracy, the Euclidean distance between the ground-truth location and the location as measured by either method is computed. For all junction types, the localization is considerably better for the method based on the recurrent long-range interaction.

3.2 Processing of Synthetic and Camera Images. In the second set of simulations, we evaluate the detection performance of the model for a variety of synthetic and real-world camera images. Receiver operator characteristic (ROC) analysis (Dayan & Abbott, 2001; Green & Swets, 1974) is used for a threshold-free evaluation of the different approaches. Details of the application of ROC analysis to junction detection are given in appendix B.

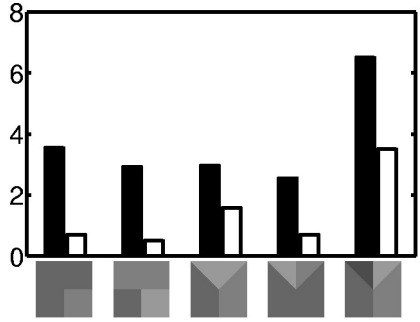

Figure 5: Distance in pixels from ground-truth location (ordinate) for L-, T-, Y-, W- and Ψ-junctions (abscissa). The deviation from ground-truth position is considerably smaller for the recurrent long-range interaction (open bars) compared to the complex cell responses (solid bars).

3.2.1 Synthetic Images. In the first simulation we employ a synthetic corner test image from Smith and Brady (1997). The ROC curve for the new method based on recurrent long-range interaction is well above the ROC curve obtained for the complex cell responses, indicating a higher accuracy of the new method (see Figure 6).

Next, we address the processing of junctions made of lines meeting at small angles. We study junctions with different numbers of lines (two, three, and four) joining at different small angles (5, 10, and 15 degrees). Input images together with the corresponding ROC curves are shown in Figure 7. For all input images, the ROC curves indicate a better performance of the junction detection based on the recurrent long-range processing.

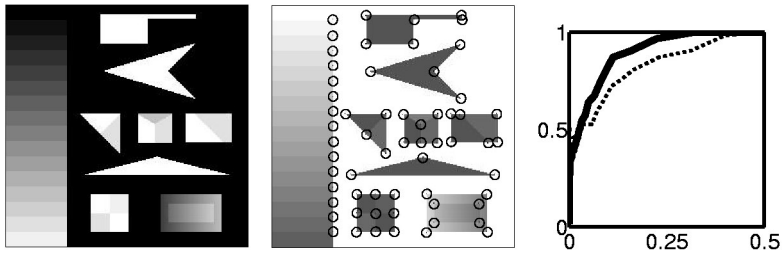

Figure 6: (Left) Synthetic test image from Smith and Brady (1997) used for the evaluation of corner detection schemes. (Middle) Overlay of ground-truth junction points. (Right) ROC curves (long-range stage, solid lines; complex cells, dashed lines). The abscissa denotes the false alarm rate, and the ordinate denotes the hit rate.

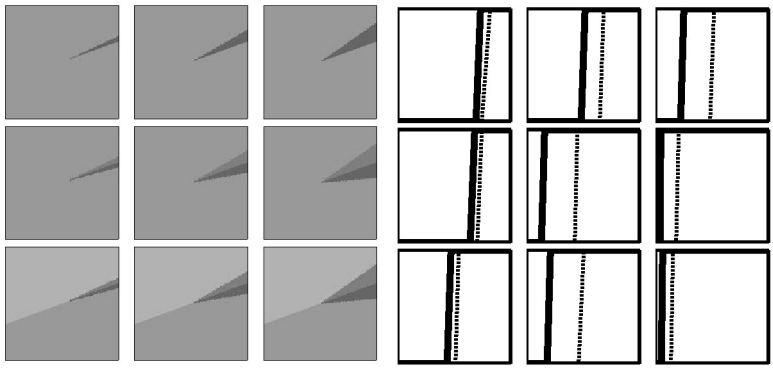

Figure 7: ROC curves for synthetic junctions of small angles. (Left panel) 3×3 arrangement of input images, each with junctions made of lines meeting at 5, 10, and 15 degrees (top to bottom row). (Right panel) ROC curves (long-range stage, solid lines; complex cells, dashed lines). The abscissa denotes the false alarm rate between zero and one, the ordinate denotes the hit rate. For better visualization, cut-outs of the left part (0 to 0.15) of the ROC curves are shown. The ROC curves indicate a higher accuracy for the results based on the recurrent long-range interaction. Note that for a single junction, the hit rate can take only two values, 0 and 1.

3.2.2 Camera Images. In this section, the junction detection performance is evaluated for various camera images. For each image, the ground-truth data are obtained by manually marking the junction points. Detection performance is evaluated for a cube image within a laboratory environment, a staircase image, a laboratory scene (Mokhtarian & Suomela, 1998), and a subimage of a plant taken from a database of natural images (van Hateren & van der Schaaf, 1998). For each image sample detection results for a fixed threshold and the full ROC curves are shown (see Figure 8).

For all images, the ROC curves obtained from the junction detection based on recurrent long-range interaction lie above the curves obtained for the feedforward complex cell responses. Thus, detection accuracy is increased by the recurrent long-range processing. The sample plots reveal two main factors that contribute to the increased detection performance. First, many false-positive responses occur based on the complex cell responses due to noisy variations of the initial orientation measurement (e.g., at the long vertical edges of the cube or at the top and bottom edges of the staircase). These variations are reduced at the long-range stage by the recurrent interaction, such that only the positions of significant orientation variations remain. Second, false-negative responses occur at the complex cell stage because localized measurements within a small aperture fail to detect junctions of locally low contrast (McDermott, 2001, 2002). Such missing responses oc-

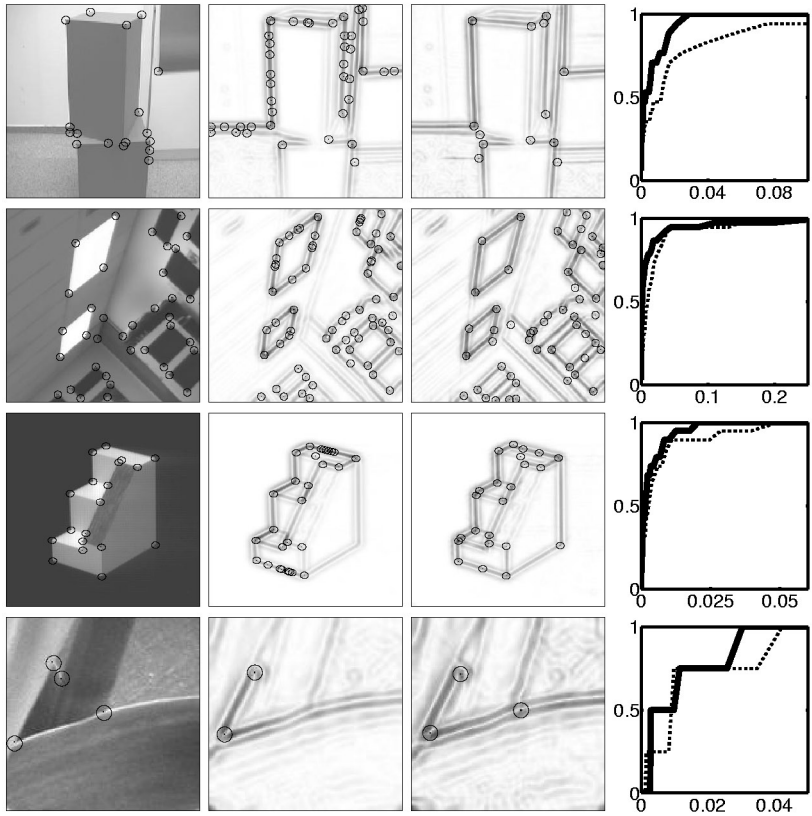

Figure 8: Simulation of the corner detection scheme for camera images. (Left to right) Input image with an overlay of manually marked ground-truth junction points; detected corners and junctions based on the complex cell responses; detected corners and junctions based on the long-range responses; and the corresponding ROC curves (long-range stage, solid lines; complex cells, dashed lines). The two gray-scale images in the second and third columns depict the edge maps based on the complex and long-range responses, respectively. These edge maps show the overall activity given by the sum across all orientations within each hypercolumn. In the edge maps, high activity is coded black, and low activity is coded white. In the ROC curves, the abscissa denotes the false alarm rate, and the ordinate denotes the hit rate. For better visualization, cutouts of the left part of the ROC curves are shown. The recurrent long-range interaction results in a decrease of false-positive and false-negative responses, leading to an increase in overall accuracy.

cur, for example, at the bottom right corner of the staircase or for the plant image. The recurrent long-range interaction enhances those small activities supported by aligned responses within a more global context, such that even junctions of low contrast can be detected.

Overall, the results obtained based on the long-range responses are superior to the results based on the purely feedforward complex cells responses. However, the results are not perfect in the sense that every corner is detected by the new method. The focus of this article is not to propose an ideal junction detector, but to show how mechanisms of recurrent long-range processing in V1 lead to a coherent representation of contours and junctions.

4 Discussion

4.1 Models of Recurrent Long-Range Interaction. A number of models have been proposed in the context of contour grouping and texture segmentation that incorporate principles of nonlocal, long-range interactions, or recurrent processing, or both (Grossberg & Mingolla, 1985b; Grossberg & Raizada, 2000; Heitger, von der Heydt, Peterhans, Rosenthaler, & Kübler, 1998; Li, 1998, 1999, 2001; Neumann & Sepp, 1999; Yen & Finkel, 1998). A comprehensive overview of these different approaches can be found in Neumann and Mingolla (2001) and Hansen (2003). These models have focused on the role of recurrent long-range processing for contour enhancement (Grossberg & Mingolla, 1985b; Li, 1998), preattentive segmentation and pop-out (Li, 1999; Yen & Finkel, 1998), recurrent interaction between two reciprocally connected cortical areas V1 and V2 (Neumann & Sepp, 1999), or the role of the laminar architecture in V1 for contrast-sensitive perceptual grouping (Grossberg & Raizada, 2000). With the exception of Heitger et al. (1998), none of these models has addressed the computation and representation of corners and junctions. The model of Heitger et al., however, relies on the explicit computation of corners and line end points by specific end-stopped filters and is purely feedforward. The idea of an explicit end-stopped filter has been challenged by physiological studies suggesting that end-stopped properties emerge from surround inhibition within a recurrent interlaminar loop (Bolz & Gilbert, 1986; Bolz, Gilbert, & Wiesel, 1989). Further, the use of a purely feedforward computation scheme neglects the role of feedback and recurrent interactions, which presumably play an important role in visual processing for figure-ground segregation (Hupé et al., 1998; Lamme, 1995; Zipser, Lamme, & Schiller, 1996).

Zucker and coworkers have investigated the process of line finding based on a relaxation labeling mechanism followed by a stage of spline approximation to minimize a global energy functional (Parent & Zucker, 1989; Zucker et al., 1989). The representation of junction configurations is only implicitly addressed there. For example, the inference of smooth contours from responses of oriented line detectors necessitates stable endings that neither shrink nor grow at discontinuities in the orientation field. By explicitly

coloring nodes of a relaxation grid and thus splitting activations into separate layers, multiple orientations can be represented at one spatial location (David & Zucker, 1990). These approaches, however, have not investigated the capability to read out junction configurations from orientation fields or analyzed the robustness of junction representation for varying opening angles. Our approach is computationally simple, requiring no explicit class labeling operation. Investigating the issue of detecting and representing corners and junctions demonstrates that our network model gives robust results for varying scene parameters. This demonstrates that such meaningful visual structure information is already available at the stage of V1.

4.2 Decoding Population Codes. In population coding, a single quantity is represented by a number or population of neurons with overlapping sensitivity or tuning curves (Oram, Földiàk, Perret, & Sengpiel, 1998; Pouget, Dayan, & Zemel, 2000). Population coding is a general principle in neural systems. Population coding is advantageous because of its robustness to both single cell damage and noisy response fluctuations (Pouget et al., 2000). The benefits of population coding of stimulus orientation within a cortical hypercolumn have been studied in a model by Vogels (1990). It has been shown that stimulus orientation can be achieved with high accuracy based on cells that have broad orientation tuning curves. Population codes also allow representing the certainty of a signal by the summed overall activity (Hinton, 1992). Further, the presence of more than one signal can be represented by a multimodal activity distribution.

Corners and junctions can be signaled if multiple activity of high amplitude occurs within a cortical hypercolumn (Zucker et al., 1989). Most schemes of decoding population codes are based on the underlying assumption that a single value is encoded (Oram et al., 1998; Seung & Sompolinsky, 1993), with a few exceptions (see Zemel, Dayan, & Pouget, 1998). To extract a single value of junction activity from the distributed, multimodal representation of orientation responses within a hypercolumn, we suggest a simple mechanism that combines two measures in a multiplicative way. The summed overall activity signals the presence or certainty of a signal (Hinton, 1992), and high circular variance signals the presence of multiple orientations. A single measurement of the sum total activity does not allow distinguishing between high activity as generated by a strong response for a single orientation or by medium responses for multiple orientations. Likewise, a single measurement of circular variance does not allow distinguishing between high activity as generated by small responses in the presence of noise or strong responses in the presence of multiple orientations. The combination of both measurements naturally allows for extracting the desired information.

4.3 Circular Variance Function. The proposed junction detection mechanism uses a measure of circular variance for the explicit representation of

distributed orientation responses within a hypercolumn. While this measure is motivated mainly computationally, circular variance is essentially a weighted sum of information locally available within a hypercolumn. We hypothesize that a related measure of junction responses can be realized by neurons in the visual cortex. First evidence comes from a study by Das and Gilbert (1999), indicating a graded specialization of neurons for the processing of corners and T-junctions. The degree of selectivity for processing corners was shown to increase with the overlap of the neuron's dendritic arborization with neighboring orientation columns. The circular variance function used in the model is invariant against rotations of the junction and increases for different numbers of lines meeting at one point. A neural circuit based on a measure of circular variance is predicted to realize a general junction detector in contrast to detectors specialized for a particular junction type such as T-junctions.

4.4 Multiscale Processing for Junction Detection. A number of studies have stressed the importance of multiple scales for the proper extraction of corner and junction information (Lindeberg, 1998; Mokhtarian & Suomela, 1998; Würtz & Lourens, 2000). The present model operates on only a single scale. However, in the recurrent interaction, the model neurons integrate information over successively increasing regions of the visual space. Thus, information of different scales is available during the temporal evolution of orientation responses. We have shown for a number of generic junction configurations that accurate, precise localization can be achieved by the proposed model without the need of tracking responses along multiple scales from coarse to fine, as suggested by Mokhtarian and Suomela (1998).

4.5 Evaluation of Junction Detection Schemes. A number of different methods have been proposed to evaluate the various approaches to corner detection. The different methods can be classified into methods based on visual inspection, localization accuracy, and theoretical analysis (Schmid et al., 2000). A simple and popular method relies on visual inspection of the detection results: A number of different detectors are applied to a set of images, and the results are presented. This method suffers from a number of drawbacks since the number of false alarms and misses as well as the precise localization may not be judged correctly. More important, results are shown only for a particular choice of the threshold separating corner from noncorner points. Such a threshold is involved in virtually every corner detection scheme, and detection results crucially depend on the proper choice of the threshold. Localization accuracy is another evaluation method and can be measured based on the correct projection of 3D scene points to 2D image points (Coelho, Heller, Mundy, Forsyth, & Zisserman, 1991; Heyden & Rohr, 1996). Since this method requires the precise knowledge of 3D points, the evaluation is restricted to simple scenes of, for example, polyhedral objects. The performance of various corner detectors can also be

assessed by theoretical analysis (Deriche & Giraudon, 1990; Rohr, 1994). Analytical studies are limited to particular configurations such as L-corners. Here we have introduced the method of ROC analysis in the context of junction detection. ROC analysis allows assessment of the capabilities of the detectors over the full range of possible thresholds for every test image. Consequently, ROC-based evaluation results are not flawed by choice of a particular threshold, which can strongly bias the obtained results.

5 Conclusions

We have proposed a novel method for corner and junction detection based on a distributed representation of orientation information within a hypercolumn (Zucker et al., 1989). The explicit representation of a number of orientations in a cortical hypercolumn is shown to constitute a powerful and flexible multipurpose scheme that can be used to code intrinsically 1D signal variations like edges as well as 2D variations like corners and junctions.

Orientation responses within a hypercolumn can be robustly and reliably computed by using contextual information. We have proposed a model of recurrent long-range interactions to compute coherent orientation responses (Hansen & Neumann, 2001). In the context of corner and junction detection, we have shown the benefits of using contextual information and recurrent interactions, leading to a considerable increase in localization accuracy and detection performance compared to a simple feedforward scheme. We have used ROC analysis for the evaluation of the different junction detection schemes. The results of ROC analysis for both synthetic and real-world camera images show that the detection performance of the new scheme is superior compared to the basic feedforward scheme. The results have been gained with a scheme that focuses on the interaction between collinear items. One might expect results to be even more impressive for a scheme that includes curves and textures. Overall, we have shown how robust and accurate junction detection can be realized based on biologically plausible mechanisms.

Appendix A: Feedforward Processing

A.1 LGN On- and Off-Cells. Responses of isotropic LGN-cells are modeled by correlation of the initial input stimulus I with a difference-of-gaussians (DoG) operator. Two types of LGN cells are modeled, on and off, which generate rectified output responses $K_{\text{on/off}}$,

$$\begin{aligned} K &= \text{DoG}_{\sigma_C,\sigma_S} \star \text{I} \\ K_{\text{on}} &= [K]^+ \\ K_{\text{off}} &= [-K]^+, \end{aligned} \tag{A.1}$$

where $\star$ is the spatial correlation operator and $[x]^+ := \max\{x, 0\}$ denotes half-wave rectification. The DoG is parameterized by the standard deviation of the center and surround gaussian ($\sigma_C = 1, \sigma_S = 3$), respectively.

A.2 Simple Cells. Simple cells in V1 have elongated subfields (on and off) that sample the input of appropriately aligned LGN responses. Input sampling is modeled by correlation with rotated, anisotropic gaussians. The gaussians are shifted perpendicular to their main axis by $\pm\tau = 3$ to model left and right subfields of an odd-symmetric simple cell. Thus, for example, for the on-channel, the equations read

$$\begin{aligned} R_{\text{on,left},\theta} &= K_{\text{on}} \star G_{\sigma_x,\sigma_y,0,-\tau,\theta} \\ R_{\text{on,right},\theta} &= K_{\text{on}} \star G_{\sigma_x,\sigma_y,0,\tau,\theta}. \end{aligned} \tag{A.2}$$

The activations of the off-channel are computed analogously.

Simple cells are modeled for two polarities (dark-light and light dark) in $O_{\max} = 4$ orientations ($\theta = 0, \pi/O_{\max}, \ldots, (O_{\max}-1)\,\pi/O_{\max}$). The standard deviations of the anisotropic gaussians are set to $\sigma_y = 1$, $\sigma_x = 3\,\sigma_y$. For each orientation, the simple cell activity is computed by pooling the two subfield responses. The equations for light-dark (ld) and dark-light (dl) simple cells read

$$\begin{aligned} S_{\text{ld},\theta} &= R_{\text{on,left},\theta} + R_{\text{off,right},\theta} \\ S_{\text{dl},\theta} &= R_{\text{off,left},\theta} + R_{\text{on,right},\theta}. \end{aligned} \tag{A.3}$$

A.3 Complex Cells. Cortical complex cells are polarity insensitive. Their response is generated by pooling simple cells of opposite polarities. Before pooling, simple cells of opposite polarities compete and are spatially blurred. The corresponding equations read

$$\begin{aligned} \widetilde{S}_{\text{ld},\theta} &= [(S_{\text{ld},\theta} - S_{\text{dl},\theta}) \star G_{\sigma_x,\sigma_y,0,0,\theta}]^+ \\ C &= \widetilde{S}_{\text{ld},\theta} + \widetilde{S}_{\text{dl},\theta}. \end{aligned} \tag{A.4}$$

Appendix B: Applying ROC for the Evaluation of Different Junction Detectors

ROC analysis allows characterizing different detectors over the full range of possible biases or thresholds. In virtually all junction detection schemes, some kind of thresholding is involved, and the detection performance crucially depends on the determination of the "optimal" threshold value. A threshold-free evaluation of different detectors as provided by ROC analysis allows separating the sensitivity of the detector from its threshold selection strategy.

ROC analysis in general is based on ground-truth verification, that is, the comparison of a detection result with ground truth. Thus, the first step to apply ROC analysis for junction detection is the specification of ground-truth junction points for each test image. For synthetic images, the ground-truth position of junction points is known from the definition of the image or can be rather easily inferred from the gray-level variations. For real-world camera images, no such ground truth exists. In this case, junction points are marked by visible inspection of an enlarged version of the image. The list of these junction points serves as an approximation of an objective ground-truth image. The next step is the estimation of junction points using a particular junction detection scheme. We include two different methods in the comparison, based on long-range responses and feedforward complex cell responses. The resulting junction responses are normalized to the range $[0; 1]$ to compensate for variations of the response amplitude across different methods. ROC curves are then computed based on the ground-truth image and the normalized junction response as follows. A threshold is varied in N steps over the full range $[0; 1]$ of junction responses, and for each value of the threshold, the proportion of true-positive (hits) and false-positive (false alarms) responses is computed. To obtain true-positive responses despite localization errors of the methods, responses are accepted within a certain error radius r_{err} around each ground-truth location. Finally, the ROC curve characterizing the detection performance of the particular method is obtained by plotting the true-positive rates against the false-positive rates.

To sum up, ROC analysis of the performance of junction detection schemes involves the following five steps:

1. Selection of an input image and determination of the ground truth-position of junction points
2. Application of a particular junction detection scheme to the image
3. Normalization of the junction responses to the range $[0; 1]$
4. Variation of a threshold in N steps from 1 to 0 and computation of the respective true-positive t_{p} and false-positive f_{p} rate
5. Plot of the ROC curve, that is, plotting t_{p} against f_{p}

The free parameters of the approach are the number of thresholds N and the error radius r_{err}. We use $N = 40$, which allows for a sufficiently fine resolution of the ROC curves. The error radius is set to $r_{\text{err}} = 3$ pixels.

Acknowledgments

We acknowledge the insightful comments and criticisms of two anonymous reviewers whose suggestions helped to improve the article considerably.

References

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science, 262*(5142), 2042–2044.

Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 339–351). Cambridge, MA: MIT Press.

Amir, Y., Harel, M., & Malach, R. (1993). Cortical hierarchy reflected in the organization of intrinsic connections in the macaque monkey visual cortex. *J. Comp. Neurol., 334*, 19–46.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychol. Rev., 61*(3), 183–193.

Biederman, I. (1985). Human image understanding: Recent research and a theory. *Computer Vision, Graphics, Image Proc., 32*(1), 29–73.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychol. Rev., 94*(2), 115–147.

Bolz, J., & Gilbert, C. D. (1986). Generation of end-inhibition in the visual cortex via interlaminar connections. *Nature, 320*(6060), 362–365.

Bolz, J., Gilbert, C. D., & Wiesel, T. N. (1989). Pharmacological analysis of cortical circuitry. *Trends Neurosci., 12*(8), 292–296.

Borg-Graham, L. J., Monier, C., & Frégnac, Y. (1998). Visual input evokes transient and strong shunting inhibition in visual cortical neurons. *Nature, 393*, 369–373.

Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci., 17*(6), 2112–2127.

Carandini, M., & Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. *Science, 264*(5163), 1333–1336.

Carpenter, G. A., & Grossberg, S. (1988). The ART of adaptive pattern recognition by a self-organizing neural network. *Computer, 21*, 77–88.

Coelho, C., Heller, A., Mundy, J. L., Forsyth, D., & Zisserman, A. (1991). An experimental evaluation of projective invariants. In *Proc. DARPA-ESPRIT Workshop on Applications of Invariants in Computer Vision, Reykjavik, Iceland* (pp. 273–293).

Das, A., & Gilbert, C. D. (1999). Topography of contextual modulations mediated by short-range interactions in primary visual cortex. *Nature, 399*, 655–661.

David, C., & Zucker, S. W. (1990). Potentials, valleys, and dynamic global coverings. *Int. J. Comput. Vision, 5*, 219–238.

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience*. Cambridge, MA: MIT Press.

DeAngelis, G. C., Freeman, R. D., & Ohzawa, I. (1994). Length and width tuning of neurons in the cat's primary visual cortex. *J. Neurophysiol., 71*(1), 347–374.

Deriche, R., & Giraudon, G. (1990). Accurate corner detection: An analytical study. In *Proc. 3rd Int. Conf. Computer Vision* (pp. 66–70). Los Alamitos, CA: IEEE Computer Society Press.

Douglas, R. J., & Martin, K. A. C. (1990). Control of neuronal output by inhibition at the axon initial segment. *Neural Comput., 2*(3), 283–292.

Fechner, G. T. (1889). *Elemente der psychophysik*. Leipzig: Breitkopf & Härtel.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field." *Vision Res., 33*(2), 173–193.

Furman, G. (1965). Comparison of models for subtractive and shunting lateral-inhibition in receptor-neuron fields. *Kybernetik, 2*, 257–274.

Gilbert, C. D., & Wiesel, T. N. (1983). Clustered intrinsic connections in cat visual cortex. *J. Neurosci., 3*, 1116–1133.

Gilbert, C. D., & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J. Neurosci., 9*(7), 2432–2442.

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics*. Huntington, NY: Krieger.

Grossberg, S. (1970). Neural pattern discrimination. *J. Theoret. Biol., 27*, 291–337.

Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychol. Rev., 92*, 173–211.

Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentation. *Percept. Psychophys.*, 38:141–171.

Grossberg, S., & Raizada, R. D. S. (2000). Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Res., 40*(10), 1413–1432.

Hansen, T. (2003). *A neural model of early vision: Contrast, contours, corners and surfaces*. Doctoral dissertation, Ulm University. Available on-line: http://vts.uni-ulm.de/query/longview.meta.asp?document_id=3022.

Hansen, T., & Neumann, H. (1999). A model of V1 visual contrast processing utilizing long-range connections and recurrent interactions. In *Proc. 9 Int. Conf. on Artificial Neural Networks (ICANN99)* (pp. 61–66). London: Institution of Electrical Engineers.

Hansen, T., & Neumann, H. (2001). Neural mechanisms for representing surface and contour features. In S. Wermter, J. Austin, & D. Willshaw (Eds.), *Emergent neural computational architectures based on neuroscience* (pp. 139–153). Berlin: Springer-Verlag.

Hansen, T., & Neumann, H. (2002). A biologically motivated scheme for robust junction detection. In H. H. Bülthoff, S.-W. Lee, T. A. Poggio, & C. Wallraven (Eds.), *Biologically motivated computer vision* (pp. 16–26). Berlin: Springer-Verlag.

Hansen, T., Sepp, W., & Neumann, H. (2001). Recurrent long-range interactions in early vision. In S. Wermter, J. Austin, & D. Willshaw (Eds.), *Emerging neural architectures based on neuroscience* (pp. 127–138). Berlin: Springer-Verlag.

Harris, C. J. (1987). Determination of ego-motion from matched points. In *Proc. Third Alvey Vision Conference* (pp. 189–192). Cambridge, UK.

Heitger, F., Rosenthaler, L., von der Heydt, R., Peterhans, E., & Kübler, O. (1992). Simulation of neural contour mechanisms: From simple to end-stopped cells. *Vision Res., 32*(5), 963–981.

Heitger, F., von der Heydt, R., Peterhans, E., Rosenthaler, L., & Kübler, O. (1998). Simulation of neural contour mechanisms: Representing anomalous contours. *Image Vision Comput., 16*, 407–421.

Heyden, A., & Rohr, K. (1996). Evaluation of corner extraction schemes using invariance methods. In *Proc. 13th Int. Conf. Pattern Recognition, Vienna, Austria* (Vol. 1, pp. 895–899). Los Alamitos, CA: IEEE Computer Society Press.

Hinton, G. E. (1992). How neural networks learn from experience. *Sci. Am., 267*(3), 144–151.

Hirsch, J. A., & Gilbert, C. D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. *J. Neurosci., 11*(6), 1800–1809.

Hodgkin, A. L. (1964). *The conduction of nervous impulses*. Liverpool: Liverpool University Press.

Holt, G. R., & Koch, C. (1997). Shunting inhibition does not have a divisive effect on firing rates. *Neural Comput., 9*(5), 1001–1013.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J. Physiol., 195*, 215–243.

Huffman, D. A. (1971). Impossible objects as nonsense sentences. In B. Meltzer & D. Michic (Eds.), *Machine intelligence 6* (pp. 295–323). Edinburgh: Edinburgh University Press.

Hupé, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature, 394*, 784–787.

Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron, 15*(4), 843–856.

Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (2000). Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *J. Neurophysiol., 84*(4), 2048–2062.

Kisvarday, Z. F., Kim, D. S., Eysel, U. T., & Bonhoeffer, T. (1994). Relationship between lateral inhibitory connections and the topography of the orientation map in cat visual cortex. *Europ. J. Neurosci., 6*(10), 1619–1632.

Lamme, V. A. F. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *J. Neurosci., 15*(2), 1605–1615.

Levine, D. S. (2000). *Introduction to neural and cognitive modeling* (2nd ed.). Mahwah, NJ: Erlbaum.

Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Comput., 10*(4), 903–940.

Li, Z. (1999). Pre-attentive segmentation in the primary visual cortex. *Spat. Vis., 13*, 25–50.

Li, Z. (2001). Computational design and nonlinear dynamics of a recurrent network model of the primary visual cortex. *Neural Comput., 13*(8), 1749–1780.

Lindeberg, T. (1998). Feature detection with automatic scale selection. *Int. J. Comput. Vision, 30*(2), 77–116.

McDermott, J. H. (2001). *Some experiments on junctions in real images*. Unpublished master's thesis, University College London. Available on-line: http://persci.mti.edu/~jmcderm.

McDermott, J. H. (2002). Psychophysics with junctions in real images. *J. Vis., 2*(7), 131a.

McLaughlin, D., Shapley, R., Shelley, M. J., & Wielaard, D. J. (2000). A neural network model of macaque primary visual cortex (V1): Orientation selectivity and dynamics in the input layer 4Cα. *Proc. Natl Acad. Sci. USA, 97*(14), 8087–8092.

Metelli, F. (1974). The perception of transparency. *Sci. Am., 230*(4), 90–98.

Michaelis, M. (1997). *Low level image processing using steerable filters*. Unpublished doctoral dissertation, Christian-Albrechts-Universität Kiel.

Mitchell, S. J., & Silver, R. A. (2003). Shunting inhibition modulates neuronal gain during synaptic excitation. *Neuron, 38*(3), 433–445.

Mokhtarian, F., & Suomela, R. (1998). Robust image corner detection through curvature scale space. *IEEE Trans. Pattern Anal. Mach. Intell., 20*(12), 1376–1381.

Mumford, D. (1991). On the computational architecture of the neocortex II: The role of cortico-cortical loops. *Biol. Cybern., 65*, 241–251.

Neumann, H., & Mingolla, E. (2001). Computational neural models of spatial integration and perceptual grouping. In T. F. Shipley & P. J. Kellman (Eds.), *From fragments to objects: Segmentation and grouping in vision* (pp. 353–400). Amsterdam: Elsevier Science.

Neumann, H., & Sepp, W. (1999). Recurrent V1–V2 interaction in early visual boundary processing. *Biol. Cybern., 81*, 425–444.

Oram, M. W., Földiàk, P., Perret, D. I., & Sengpiel, F. (1998). The "ideal homunculus": Decoding neural population signals. *Trends Neurosci., 21*(6), 259–265.

Parent, P., & Zucker, S. W. (1989). Trace inference, curvature consistency, and curve detection. *IEEE Trans. Pattern Anal. Mach. Intell., 11*(8), 823–839.

Parida, L., & Geiger, D. (1998). Junctions: Detection, classification, and reconstruction. *IEEE Trans. Pattern Anal. Mach. Intell., 20*(7), 687–698.

Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *J. Neurophysiol., 82*(5), 2490–2502.

Pasupathy, A., & Connor, C. E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *J. Neurophysiol., 86*(5), 2505–2519.

Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Res., 33*(7), 993–999.

Polat, U., & Sagi, D. (1994). The architecture of perceptual spatial interactions. *Vision Res., 34*(1), 73–78.

Pouget, A., Dayan, P., & Zemel, R. (2000). Information processing with population codes. *Nat. Rev. Neurosci., 1*(2), 125–132.

Prescott, S. A., & Koninck, Y. D. (2003). Gain control of firing rate by shunting inhibition: Roles of synaptic noise and dendritic saturation. *Proc. Natl Acad. Sci. USA, 100*(4), 2076–2081.

Pugh, M. C., Ringach, D. L., Shapley, R., & Shelley, M. J. (2000). Computational modeling of orientation tuning dynamics in monkey primary visual cortex. *J. Comput. Neurosci., 8*(2), 143–159.

Ringach, D. L., Hawken, M. J., & Shapley, R. (1997). Dynamics of orientation tuning in macaque primary visual cortex. *Nature, 387*, 281–284.

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

Rohr, K. (1994). Localization properties of direct corner detectors. *J. Math. Imag. Vision, 4*, 139–150.

Rubin, N. (2001). The role of junctions in surface completion and contour matching. *Perception, 30*(3), 339–366.

Salin, P.-A., & Bullier, J. (1995). Corticocortical connections in the visual system: Structure and function. *Physiol. Rev., 75*(1), 107–154.

Sandell, J. H., & Schiller, P. H. (1982). Effect of cooling area 18 on striate cortex cells in the squirrel monkey. *J. Neurophysiol., 48*(1), 38–48.

Schmid, C., Mohr, R., & Bauckhage, C. (2000). Evaluation of interest point detectors. *Int. J. Comput. Vision, 37*(2), 151–172.

Schmidt, K. E., Goebel, R., Löwel, S., & Singer, W. (1997). The perceptual grouping criterion of collinearity is reflected by anisotropies of connections in the primary visual cortex. *Europ. J. Neurosci., 9*, 1083–1089.

Seung, H. S., & Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proc. Natl Acad. Sci. USA, 90*(22), 10749–10753.

Sinha, P., & Adelson, E. H. (1993). Recovering reflectance in a world of painted polyhedra. In *Proc. Fourth International Conference on Computer Vision (ICCV' 93), Berlin, Germany* (pp. 156–163). New York: IEEE.

Smith, S., & Brady, J. (1997). SUSAN—a new approach to low level image processing. *Int. J. Comput. Vision, 23*(1), 45–78.

Todorović, D. (1997). Lightness and junctions. *Perception, 26*(4), 379–394.

van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. R. Soc. London (B), 265*(1394), 359–366.

Vogels, R. (1990). Population coding of stimulus orientation by striate cortical cells. *Biol. Cybern., 64*, 25–31.

Weber, E. H. (1905). Tastsinn und Gemeingefühl. *Handwörterbuch der Physiologie*. Leipzig: Engelmann.

Würtz, R. P., & Lourens, T. (2000). Corner detection in color images through a multiscale combination of end-stopped cortical cells. *Image Vision Comput., 18*(6–7), 531–541.

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

Zemel, R. S., Dayan, P., & Pouget, A. (1998). Probabilistic interpretation of population codes. *Neural Comput., 10*(2), 403–430.

Zipser, K., Lamme, V. A. F., & Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *J. Neurosci., 16*(22), 7376–7389.

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

Received July 29, 2002; accepted October 7, 2003.