# Neural Mechanisms for Representing Surface and Contour Features

Thorsten Hansen and Heiko Neumann

Universität Ulm, Abt. Neuroinformatik, D-89069 Ulm, Germany
(hansen,hneumann)@neuro.informatik.uni-ulm.de

**Abstract.** Contours and surfaces are basic qualities which are processed by the visual system to aid the successful behavior of autonomous beings within the environment. There is increasing evidence that the two modalities of contours and surfaces are processed in separate, but interacting visual streams or sub-systems. Neurons at early stages in the visual system show strong responses only at locations of high contrast, such as edges, but only weak responses within homogeneous regions. Thus, for the processing and representation of surfaces, the visual system has to integrate sparse local measurements into a dense, coherent representation. We suggest a mechanism of confidence-based filling-in, where a confidence measure ensures a robust selection of sparse contrast signals. The new mechanism supports the generation of surface representations which are invariant against size and shape transformation. The filling-in process is controlled by contour or boundary signals which stop the filling-in of contrast signals at region boundaries. Localized responses to contours are most often noisy and fragmented. We suggest a recurrent processing scheme for the extraction of contours that incorporates long-range connections. The recurrent long-range processing enhances coaligned activity which is consistent within a more global context, while inconsistent noisy activity is suppressed. The capability of the model is shown for noisy synthesized and natural stimuli.

## 1 Introduction

Experimental studies indicate the existence of distinct perceptual subsystems in human vision, one that is concerned with *contour extraction* and another that assigns *surface properties* to bounded regions. The emerging picture from the experimental investigations is one in which shape outlines are initially extracted, followed by the assignment of attributes such as texture, color, lightness, brightness or transparency to regions [10,38,6,17]. Several perceptual completion phenomena [35] suggest that, on a functional level, regions inherit local border contrast information by means of "spreading mechanisms" or "filling-in" [32,7]. The assignment of surface properties would then be dependent on the determination of stimulus contrast in various feature dimensions, such as luminance, motion direction and velocity, and depth, that would be used to fill-in bounded regions.

S. Wermter et al. (Eds.): Emergent Neural Computational Architectures, LNAI 2036, pp. 139–153, 2001.

## 2 Neural Mechanisms for Representing Surface Features

The problem of deriving a dense representation of surface quality, such as brightness or color, from local estimates, such as luminance or chromatic border contrast, is inherently ill-posed: there exists no unique solution nor is the solution guaranteed to be stable. Such an inverse problem needs to be regularized in the sense that certain constraints have to be imposed on the space of possible solutions. The constraint of generating a smooth surface, as formalized by minimizing the first order derivatives, leads to a linear diffusion process with a simple reaction term [30].

In filling-in theory, feature signals which provide the source term of the filling-in process are modeled as cells with circular receptive fields (RFs) such as retinal ganglion cells or LGN cells. In previous filling-in models, such cells are modeled to exhibit strong responses even to homogeneous regions [8,18]. Physiological studies however show that retinal ganglion cells respond strongly only at positions of luminance differences or contrasts [11]. Motivated by these results we use sparse contrast signals with no response to homogeneous regions. The sparse nature of signals necessitates additional confidence signals for the filling-in process. Confidence signals indicate the positions of valid contrast response to be taken as source for the filling-in process. Having established the link between models of perceptual data for biological vision and the mathematical frameworks of regularization theory this lead to the proposal of *confidence-based filling-in* [30].

### 2.1 Evidence for Neural Filling-in

Filling-in models are based on the assumption of distributed, topographically organized maps of boundaries and regions [25]. This assumption has been questioned in favor of a non-topographic or sparse coding of contrasts and boundaries using a compact symbolic code or a sparse localized code [37]. Regarding filling-in, both visual scientists and philosophers have argued against the logical need for a neural spreading of activity and against a continuous representation of the brightness profiles (for review, see [35]). It has been suggested that instead of filling-in, the brain could simply assign a symbolic brightness label to a bounded region or could ignore the absence of direct neural support. In support of the filling-in hypothesis, there is ample empirical evidence, mostly from psychophysics, that brightness perception indeed depends on a neural activity spreading.

Evidence comes from a study where a visual masking paradigm is used to investigate two issues [32]. First, the role of edge information in determining the brightness of homogeneous regions, and second the temporal dynamics of brightness perception. If brightness perception relies on some form of activity spreading, it should be possible to interrupt this spreading process. In the experiment, a target of a bright disk is followed by a mask (e.g., a smaller circle or a C-shape), which is presented at variable time intervals. For an interstimulus interval of about 50–100 ms, the brightness of the central area is highly dependent on the shape of the mask. For example, for a C-shaped mask, a darkening

of the middle region is observed, with the bright region "protruding" inside the C. For a circular-shaped mask, an inner dark disk is perceived. Both these results are consistent with the hypothesis that brightness signals are generated at the borders of their target stimuli and propagate inward. Furthermore, it has been demonstrated that for larger stimuli maximal suppression occurs later. This finding supports the view that filling-in is an active spreading of neural activity, i.e., a process which takes time.

Recently, a similar masking paradigm has been used to investigate brightness filling-in within texture patterns [7]. Again the spreading could be blocked by the mask if the interstimulus interval is in accordance with the propagation rate required to travel the distance between boundary and mask position. Results of a study employing Craik-O'Brien-Cornsweet (COC) gratings point in the same direction: For higher spatial frequencies of the grating (i.e., for smaller distances) the effect was stronger and persisted to higher temporal frequencies of COC contrast reversal [9,34].

In summary, these studies are suggestive of active neural filling-in processes that are initiated at region edges. Using brightness filling-in, the brain generates a spatially organized representation through a continuous propagation of signals, a process that takes time [35,30].

### 2.2 Mathematical Models for Filling-in

To introduce concepts, we consider the task of generating a continuous representation of surface layout as one of painting or coloring an empty region [26]. The task thus consists of generating an internal representation of surface properties from given data. Individual surfaces occur at different sizes and with various shapes. Therefore, any such mechanism has to be insensitive to such size and shape variations.

Models of brightness perception were among the first to explore the dichotomy of boundary and surface subsystems. Based on stabilized image studies, it has been proposed that the perception of brightness can be modeled by filling-in processes. Filling-in models suggest that feature measures are used in the determination of surface appearance through a process of lateral spreading, or diffusion [12]. The basic ideas are formalized in a model of complementary boundary and surface systems (Boundary Contour System/Feature Contour System, BCS/FCS) [8,17]. In a nutshell, BCS/FCS processing occurs as follows: The BCS extracts boundaries via a hierarchy of processing levels, defining a segmentation of the initial input image into compartmental boundaries. Within the FCS, these boundaries control the lateral spreading or diffusion of contrast-sensitive input signals. This proposal qualitatively accounts for a white variety of brightness phenomena, including, e.g., simultaneous contrast, brightness assimilation and the COC effect [18]. An extension of the model accounts for trapezoidal and triangular Mach bands, low- and high-contrast missing fundamental stimuli and sinusoidal waves, among others [33].

**Standard Filling-in Equation.** The filling-in equation that is used in several models of early vision [8,18,14,33,16] is equivalent to a linear inhomogeneous diffusion with reaction term [30]. The reaction term consists of contrast-sensitive input signals $K$ and a passive decay of activity (with rate $\alpha$). The diffusion term describes the nearest-neighbor coupling $\mathcal{N}_i$ of filling-in activities which is locally controlled by permeability signals $P$ (inhomogeneous diffusion). Permeability signals are a monotonically decreasing function of boundary or contour signals, i.e., high contour signals imply low permeability and vice versa. In all, the discretized equation for filling-in activity $U$ reads

$$\partial_t U_i = \underbrace{-\alpha U_i + K_i}_{\text{reaction term}} + \underbrace{\sum_{j \in \mathcal{N}_i} (U_j - U_i)\, P_{ij}}_{\text{diffusion term}} \ , \tag{1}$$

where $\partial_t$ denotes partial differentiation with respect to $t$. Discrete spatial locations are denoted by $i$ and $j$. The nearest neighbor coupling is given by $\mathcal{N}_i = \{i-1, i+1\}$ for the 1D case and $\mathcal{N}_{ij} = \{(i-1, j), (i+1, j), (j-1, i), (j+1, i)\}$ for the 2D case.

**Confidence-based Filling-in Equation.** Previous models of filling-in use a dense representation of contrast-sensitive feature signals as source for the filling-in process. Cells at early stages of the visual system, such as retinal ganglion cells, show strong responses only at luminance discontinuities. Given the sparseness of contrast signals which are zero within homogeneous regions, the visual system has to compute a dense brightness surface from local contrast estimates. Such inverse problems are generally ill-posed in the sense of Hadamard [41,36, 2]. This means that the existence and uniqueness of a solution and its continuous dependence on the data cannot be guaranteed since the measurements are sparse and may be noisy. The solution to the problem has to be regularized such that proper constraints are imposed on the function space of solutions. Such a constraint is the *smoothness* of the solution, for example. Smoothness can be characterized by minimizing the first order derivatives of the desired solution. The goal is to minimize both the local differences between the measured data and the reconstructed function values (data term) and the stabilizing functional imposed on the function (smoothness term). Minimizing a quadratic functional finally leads to the discretized version of a new filling-in equation, where an additional *confidence signal* $Z$ steers the contribution of the data term [30]:

$$\partial_t U_i = (-\alpha U_i + K_i)\, Z_i + \sum_{j \in \mathcal{N}_i} (U_j - U_i)\, P_{ij} \ . \tag{2}$$

Note that for constant unit-valued confidence signals $Z = 1$ confidence-based filling-in (Eq. 2) is equivalent to standard filling-in (Eq. 1).

Confidence signals are in the range $[0; 1]$. Zero confidence signals indicate positions where no data are available, while unit-valued confidence signals occur at region boundaries and signal positions of reliable contrast measurements.

Consequently, we suggest that an intermediate representation in the processing of contour signals, namely complex cell responses $C$, are involved in the computation of confidence signals. A candidate mechanism is

$$Z_i = \beta\, C_i + \varepsilon \ , \tag{3}$$

where $\beta$ is a scaling parameter, and $\varepsilon$ is a small tonic input to achieve well-posedness of the filling-in process. It is suggested that the complex cell interaction incorporates the self-normalizing properties of a shunting interaction [22], to generate signals in a bounded range such as $[0; 1]$.

For a detailed description of the model equations and the parameters used the reader is referred to [30].

### 2.3 Simulation Results

In this section we present simulation result which show that the proposed scheme of confidence-based filling-in exhibits basic properties which makes it suitable for the computation of surface properties in early vision. Results are compared for confidence-based filling-in and the corresponding standard filling-in by setting $Z = 1$.

First, we demonstrate the independence of the brightness predictions of confidence-based filling-in on the shape and size of the regions (Fig. 1 and 2). The mechanism of confidence-based filling-in is then applied to psychophysical stimuli (Fig. 3). In order to demonstrate the model's capacity to deal with real world data, we finally show results of processing real camera images (Fig. 4).

**Invariance Properties.** The first investigation focuses on the properties of the filling-in mechanisms and their dependency on the parameter settings and the size of the region to be filled-in. We start with a simple luminance pattern that shows a light square on a dark background (Fig. 1). The brightness signals generated by the standard filling-in mechanism tend to bow depending on the strength of the permeability coefficient. An increase in the permeability helps

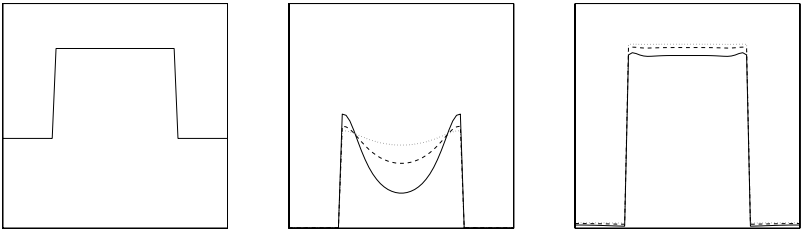

**Fig. 1.** Generation of brightness appearance for a rectangular test pattern utilizing mechanisms of standard and confidence-based filling-in. *Left to right:* Luminance profile, simulation results for standard filling-in and confidence-based filling-in under variations of the permeability parameter

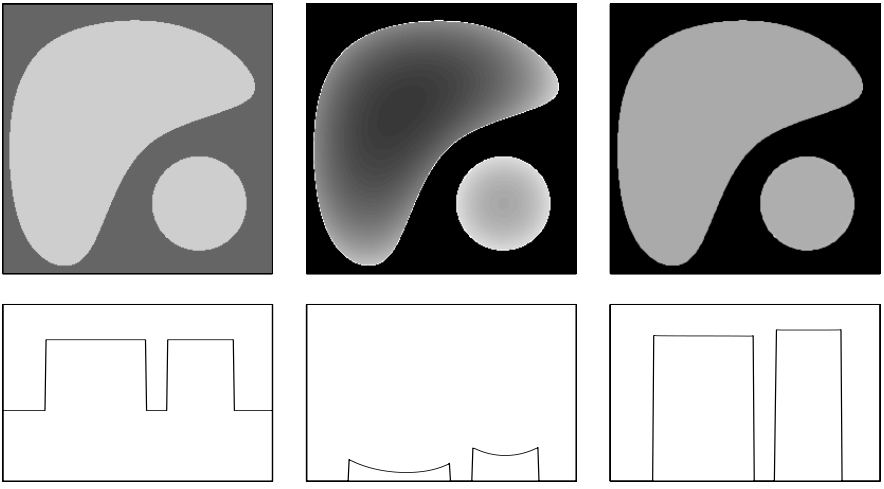

**Fig. 2.** Filled-in brightness signals for shapes of different size but same luminance level. Signal representations are generated by the filling-in mechanisms using the parameter settings that achieved proper results for the square test pattern in Fig. 1. *Top row, left to right:* Input luminance pattern, brightness signal generated by standard filling-in and by confidence-based filling-in. *Bottom row:* Corresponding profiles of the luminance function and the brightness patterns taken along the 2D picture diagonals (from upper left to lower right corner)

generating flat signals (Fig. 1, middle). The corresponding brightness patterns generated by confidence-based filling-in remain invariant against these parameter changes and are always flat (Fig. 1, right). Next, the same mechanisms have been applied to another test image that contains shapes of different form and size but the same luminance level. The results reveal the potential weaknesses of standard filling-in: Depending on the size or diameter of a pattern, which is unknown, the brightness signals appear at a different amplitude and show different amounts of bowing (Fig. 2, middle). With the confidence-based filling-in mechanism the brightness patterns appear homogeneous and of almost the same brightness (Fig. 2, right). We conclude that confidence-based filling-in helps to generate a brightness representation that is largely invariant against shape and size transformations, thus improving the robustness of filling-in mechanisms.

**Psychophysical Data on Brightness Perception.** In this section we demonstrate the ability of confidence-based filling-in to process classical luminance patterns that have been investigated in brightness perception. We particularly focus on remote border contrast effects and their creation of brightness differences. These cases provide examples of the crucial role of edges in determining the brightness appearance. For example, two regions of equal uniform luminance separated by a "cusp edge" appear differently bright, the so-called Craik-O'Brien-Cornsweet (COC) effect. These types of stimuli have been identified

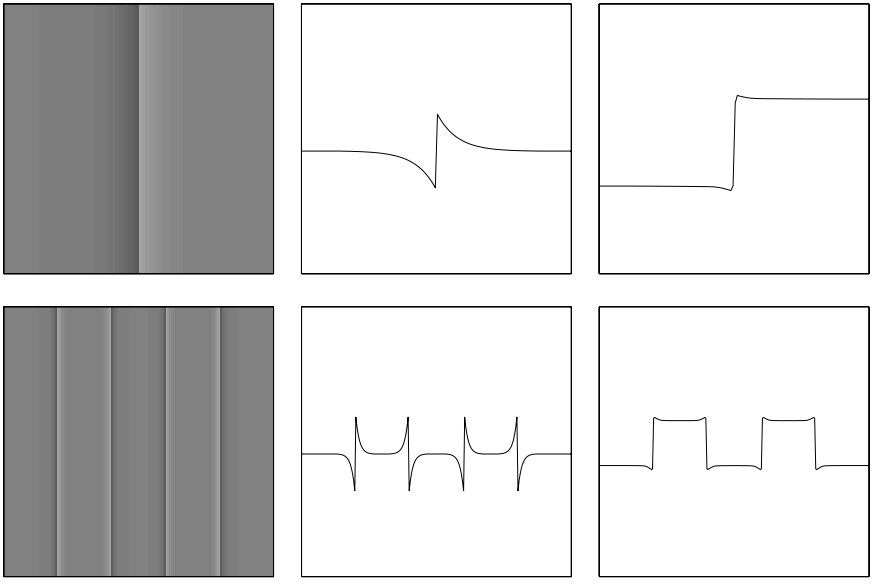

**Fig. 3.** Filled-in brightness signals for a standard COC stimulus and a COC grating *(bottom row)* made of cusps of opposite contrast polarity. *Left to right:* Input luminance pattern with corresponding profile, and the profile of the brightness pattern generated by confidence-based filling-in

as the most challenging ones for alternative theories of brightness perception, such as, for example, filter theories. In fact, yet, only filling-in models appear to properly predict the brightness appearance for COC stimuli and their variants (compare [4]). The processing of a standard COC stimulus is shown in Fig. 3 (top row). A COC stimulus consist of a cusp edge, separating two regions of equal luminance. Both regions seem to be of different brightness, where the region which is associated with the negative lobe of the cusp is perceived as a uniformly darker region compared to the right region. Confidence-based filling-in correctly predicts this effect (Fig. 3, top right), as do previous filling-in models.

A COC grating (Fig. 3, bottom row) consists of a sequence of cusp edges having pairwise opposite contrast polarities. This stimulus is perceived as a series of alternating dark and bright stripes similar to a square wave. The temporal dynamics of brightness perception in such COC arrangement is consistent with a filling-in mechanism. Confidence-based filling-in, at equilibrium, correctly predicts the appearance of the final brightness square wave pattern (Fig. 3, bottom right).

**Real World Application.** In order to demonstrate the functional significance of the proposed mechanism, we show the processing results for a camera image of a real object. In order to exclude any possible influences from 3D effects, e.g.,

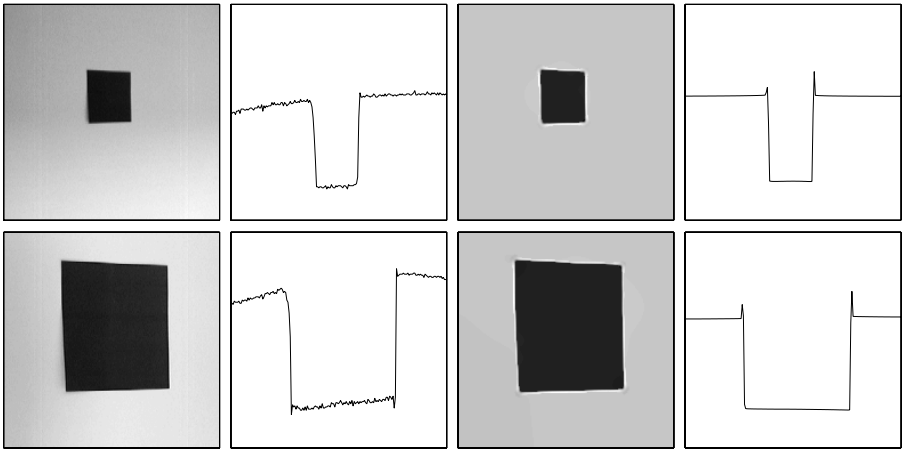

**Fig. 4.** Processing the camera images of a flat 3D object acquired from different distances. The target object of low reflectance is attached on a lighter background surface and illuminated by a primary light source that generates a visible illumination gradient. *Top:* Input intensity image of the object at a larger viewing distance and a profile section *(left pair)* together with the corresponding filling-in result and a profile section *(right pair). Bottom:* Corresponding input representations and processing results for the object at a closer viewing distance

by shadowing or variations in surface orientations, we used a card-board that has been attached to a flat background surface. This intrinsically flat scene was directly illuminated by a point-like light source at a distance of approximately 2 m. This generates a significant intensity gradient in the original intensity image. The target surface has been imaged from two different distances at about 2 and 1 m, respectively.

Simulation results show that the mechanism of confidence-based filling-in is capable of generating a representation of homogeneous surface properties (Fig. 4). The result is independent of the projected region size, thus showing the property of size invariance. Also the illumination gradient is discounted and the noise is successfully suppressed.

### 2.4 Outlook

The proper restoration of reference levels remains a deficit of filling-in functionality. The use of DC-free contrast signals discounts the illuminant, but at the same time destroys all information about the reference levels of contrast signals. Several approaches have been advocated to solve this problem, such as directional filling-in [1] or an extra luminance channel [18,28], but fail to discriminate, e.g., COC stairs from luminance stairs, or are flawed by missing physiological evidence. We suggest that a multi-scale approach [40] together with the localized coding of luminance information at contrast positions may solve the problem.

## 3 Contour Processing

The generation of brightness representations by means of filling-in relies on the proper computation of contour signals. For the filling-in process, robust, reliable contour extraction is important, since contour signals are used to determine permeability signals which control the lateral spreading of activities (cf. Eqns. 1 and 2). Contour signals must not suffer from high amplitude variations to allow for a stable representation of brightness surfaces. Initial contrast measurements, however, which define the first processing stage in the computation of contour signals, are often noisy and fragmented. Therefore, an important task in early visual processing is to determine the salient or prominent contours out of an array of noisy, cluttered contrast responses.

How can this task be accomplished? We suggest a computational framework involving *long-range connections*, *feedback*, and *recurrent interactions*. The task of contour extraction cannot be solved solely on the basis on the incoming data alone, but requests for additional constraints and assumptions on the shape of frequently occurring contours. An important principle of salient contours is that they obey the Gestalt law of good continuation. It has been suggested that horizontal *long-range connections* found in the superficial layers of early visual areas like V1 and V2 provide a neural implementation of the law of good continuation [39]. The assumptions or a priori information such as expressed in the law of good continuation have to be carefully matched against the incoming data. We suggest that *feedback* plays a central role in this matching process by selectively enhancing those feedforward input signal which are consistent with the assumptions. The interaction between feedforward data and feedback assumptions requires certain time steps. In each step the result of the interactions is recursively fed into the same matching process. Such a process of *recurrent interaction* might be used by the brain to determine the most stable and consistent representation depending on both the assumptions and the given input data.

Motivated by empirical findings we present a model of *recurrent long-range interaction in the primary visual cortex* for contour processing.

### 3.1 Computational Model

The computational model incorporates feedforward and feedback processing, lateral competitive interaction and horizontal long-range integration, and localized receptive fields for oriented contrast processing. The model architecture is defined by a sequence of preprocessing stages and a recurrent loop based on long-range interaction. The model realizes a simplified architecture of V1 [13] and is outlined in Fig. 5. The computational role of the proposed circuit is to enhance the salient contours and to suppress noisy activities. The circuit compensates for variations of amplitude strength and orientation selectivity in the initial contrast measurements along the contour. This property allows for the robust computation of closed contours to be used in the filling-in process.

The model uses *modulating* feedback, i.e., initial bottom-up activity is necessary to generate activity. The model of V1 thus does not allow for the creation

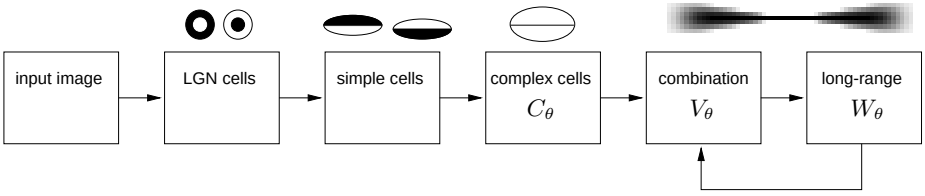


**Fig. 5.** Overview of the model stages together with a sketch of the sample receptive fields of cells at each stage for 0° orientation. For the long-range stage, the spatial weighting function of the bipole filter is shown

of illusory contours. Illusory contours evoke cell responses in V2 [42] and have been investigated in a model of V1–V2 interactions [31].

We propose a functional architecture for recurrent processing. In this architecture of two interacting regions, let them be cortical layers or areas, each region has a distinctive purpose. 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 [31,21].

In the *feedforward path*, the initial luminance distribution is processed by isotropic LGN-cells, orientation-selective simple and complex cells. 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. A more elaborated processing in the feedforward path would make use of, e.g., nonlinear processing at the level of LGN cells and simple cells [19,29]. The computation in the feedforward path is detailed in [20]. In our model, complex cell responses $C_\theta$ as output of the feedforward path (cf. Fig. 5) provide an initial local estimate of contour strength, position and orientation which is used as bottom-up input for the recurrent loop. The *recurrent loop* has two stages, a combination stage where bottom-up and top-down inputs are integrated, and a stage of long-range interaction. At the *combination* stage, feedforward inputs $C_\theta$ and feedback inputs $W_\theta$ are added and subject to shunting interaction

$$\mathrm{net}_\theta = C_\theta + \delta_V W_\theta \ , \tag{4}$$

$$\partial_t V_\theta = -\alpha_V V_\theta - \beta_V V_\theta \, \mathrm{net}_\theta + \mathrm{net}_\theta \ . \tag{5}$$

The equation is solved at equilibrium, resulting in a normalization of activity

$$V_\theta = \beta_V \frac{\mathrm{net}_\theta}{\alpha_V + \mathrm{net}_\theta} \ . \tag{6}$$

The weighting parameter $\delta_V = 2$ is chosen so that dimensions of $C_\theta$ and $W_\theta$ are approximately equal, the decay parameter $\alpha_V = 0.2$ is chosen small compared to $\mathrm{net}_\theta$ and $\beta_V = 10$ scales the activity to be sufficiently large for the subsequent long-range interaction. For the first iteration step, feedback responses $W_\theta$ are set to $C_\theta$.

At the *long-range* stage, the contextual influences on cell responses are modeled. Directional sensitive long-range connections provide the excitatory input. The inhibitory input is given by undirected interactions in both the spatial and orientational domain. Long-range connections are modeled by a bipole filter [17]. The spatial weighting function of the bipole filter is narrowly tuned to the preferred orientation, reflecting the highly significant anisotropies of long-range fibers in visual cortex [39,5] (see Fig. 5, top right). The size of the bipole is about twice the size of the RF of a complex cell.

Essentially, excitatory input is provided by correlation of the feedforward input with the bipole filter $B_\theta$. A cross-orientation inhibition prevents the integration of cells responses at positions where orthogonal responses also exists. The excitatory input is governed by

$$\mathrm{net}_\theta^+ = \left[V_\theta - V_{\theta_\perp}\right]^+ \star B_\theta \ , \tag{7}$$

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

The profile of the bipole filter is defined by a directional term $D_\varphi$ and a proximity term generated by an isotropically blurred circle $C_r \star G_\sigma$ where $r = 25$, $\sigma = 3$. The detailed equations read

$$B_{\theta,\alpha,r,\sigma}(x, y) = D_\varphi \cdot C_r \star G_\sigma \tag{8}$$

$$D_\varphi = \begin{cases} \cos(\frac{\pi/2}{\alpha}\varphi) & \text{if } \varphi < \alpha \\ 0 & \text{otherwise} \end{cases} , \tag{9}$$

where $\varphi$ is defined as $\mathrm{atan2}\,(|y_\theta|, |x_\theta|)$ and $(x_\theta, y_\theta)^\mathsf{T}$ denotes the vector $(x, y)^\mathsf{T}$ rotated by $\theta$. The parameter $\alpha = 10°$ defines the opening angle of $2\alpha$ of the bipole. The factor $\frac{\pi/2}{\alpha}$ maps the angle $\varphi$ in the range $[-\alpha; \alpha]$ to the domain $[-\pi/2; \pi/2]$ of the cosine function with positive range.

Responses which are not salient in the sense that nearby cells of similar orientation preference also show strong activity should be diminished. Thus an inhibitory term is introduced which samples activity from both orientational $\widetilde{g}_{\sigma_o,\theta}$, $\sigma_o = 0.5$, and spatial neighborhood $G_{\sigma_{\mathrm{sur}}}$, $\sigma_{\mathrm{sur}} = 8$,

$$\mathrm{net}_\theta^- = \mathrm{net}_\theta^+ \circledast \widetilde{g}_{\sigma_o,\theta} \star G_{\sigma_{\mathrm{sur}}} \ , \tag{10}$$

where $\circledast$ denotes correlation in the orientation domain. The orientational weighting function $\widetilde{g}_{\sigma_o,\theta}$ is implemented by a 1D Gaussian $g_{\sigma_o}$, discretized on a zero-centered grid of size $o_{\max}$, normalized, and circularly shifted so that the maximum value is at the position corresponding to $\theta$. The parameterization of the spatial inhibitory neighborhood results in an effective spatial extension of about half the size of the bipole filter.

Excitatory and inhibitory term combine through shunting interaction

$$\partial_t W_\theta = -\alpha_W W_\theta - \eta^- W_\theta \,\mathrm{net}_\theta^- + \beta_W V_\theta \left(1 + \eta^+ \,\mathrm{net}_\theta^+\right) \ . \tag{11}$$

The equation is solved at equilibrium, resulting in a divisive interaction

$$W_\theta = \beta_W \frac{V_\theta \left(1 + \eta^+ \,\mathrm{net}_\theta^+\right)}{\alpha_W + \eta^- \,\mathrm{net}_\theta^-} \ . \tag{12}$$

where $\eta^+ = 5$, $\eta^- = 2$ and $\beta_W = 0.001$ are scale factors and $\alpha_W = 0.2$ is a decay parameter. The multiplicative contribution of $V_\theta$ ensures that long-range connections have a modulating rather than generating effect on cell activities [23, 24]. The result of the long-range stage is fed back and combined with the feedforward complex cell responses, thus closing the recurrent loop. The shunting interactions ensure a saturation of activities after a few recurrent cycles.

### 3.2 Simulation Results

In a first simulation a synthetic stimulus of a noisy square is employed. Figure 6 demonstrates the functionality of lateral long-range interaction for the enhancement of coherent structure. Outline contrasts are detected and subsequently enhanced such that the activities of salient contrast as well as orientation significance is optimized. Figure 7 shows the results of processing an image of a laboratory scene. Initial complex cell activations generated for localized high contrast contours are further stabilized. Initially weak activations in coherent spatial arrangements are enhanced. Spatial locations where high amplitude contrast responses exist in multiple orientation channels indicate the presence of corners and junctions. The results demonstrate that noisy low contrast arrangements can be significantly enhanced to form elementary items of smooth contour segments. Beyond the enhancement of coherent contours, the proposed scheme is able to enhance contour responses at corner and junction configurations. These higher order features play a significant role in object recognition and depth segregation (e.g., [3]).

## 4 Summary

We have presented a computational framework for the processing of discontinuities and homogeneous surface properties.

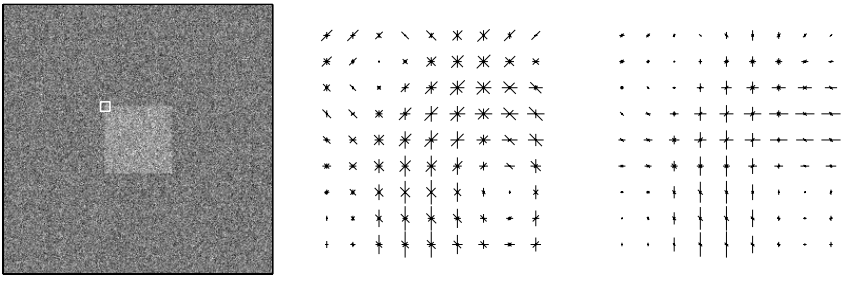

**Fig. 6.** Processing of a square pattern with additive high amplitude noise. *Left to right:* Input image and close-up of the upper left corner (white square inset in the input image) for complex cell responses and long-range responses. In the close-ups, three important properties of the long-range interaction can be seen: i) enhancement of the orientation coaligned to the contour, ii) suppression of noisy activity in the background, and iii) preservation of the significant orientations at corners

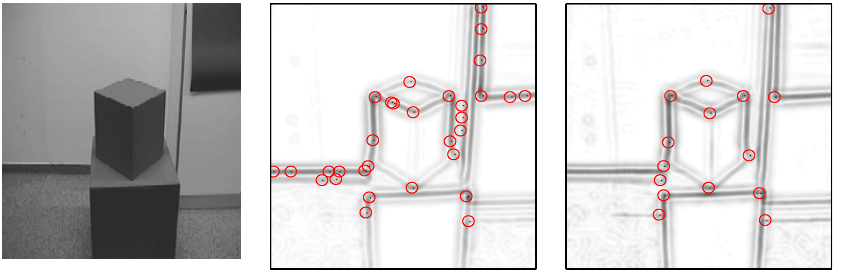

**Fig. 7.** Enhancement of activity distribution and detection of corner and junction features in a laboratory scene. *Left to right:* Input image, complex cell responses, and long-range responses. Locations of corners and junctions are marked with circles and indicate positions with significant responses in more than one orientation channel. At the complex cell stage, many false responses are detected due to noisy variations of the initial orientation measurement. Such variations have been reduced at the long-range stage, and only the positions of significant variations at corners are signaled

*Surface properties* such as brightness can be computed from sparse contrast data. Confidence signals are used to discriminate positions of reliable measurements of contrast data from positions where no data is available. The suggested mechanism of confidence-based filling-in allows to generate a size invariant brightness representation even on the basis of sparse input data. Furthermore, perceptual phenomena and real world applications are successfully processed.

For the filling-in process, the proper extraction of *contours* is important. For the processing of *discontinuities* such as contours and junctions, we have suggested a framework of recurrent interaction, using feature integration by long-range connections to evaluate feedforward signals within a broader context. Modulating feedback then selectively enhances those features which fit into the context. The suggested circuit of long-range interactions is an instantiation in the domain of early vision of this general scheme. We show that a single circuit is sufficient to solve basic tasks in early vision, such as contour enhancement, noise suppression and corner enhancement.

While the importance of contour signals for various tasks, such as object recognition, is generally acknowledged, the need for an explicit and intrinsically redundant representation of extended brightness regions is subject to intense debate. Whether such a representation is crucially involved in conscious human brightness perception or is helpful for behavioral tasks such as grasping or object recognition of occluded objects [15,27] is a challenging question to be answered by future research. Models of surface completion are helpful by integrating empirical results into a precise computational and algorithmic description.

## References

1. K. F. Arrington. Directional filling-in. *Neural Comput.*, 8:300–318, 1996.
2. M. Bertero, T. Poggio, and V. Torre. Ill-posed problems in early vision. *Proc. IEEE*, 76(8):869–889, 1988.

3. I. Biederman. Human image understanding: Recent research and a theory. *CVGIP*, 32(1):29–73, 1985.
4. B. Blakeslee and M. E. McCourt. A multiscale spatial filtering account of the white effect, simultaneous brightness contrast and grating induction. *Vision Res.*, 38:4361–4377, 1999.
5. W. H. Bosking, Y. Zhang, B. Schofield, and D. Fitzpatrick. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci.*, 17(6):2112–2127, 1997.
6. P. Bressan, E. Mingolla, L. Spillmann, and T. Watanabe. Neon color spreading: A review. *Perception*, 26:1353–1366, 1997.
7. G. Caputo. Texture brightness filling-in. *Vision Res.*, 38(6):841–851, 1998.
8. M. Cohen and S. Grossberg. Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Percept. Psychophys.*, 36:428–456, 1984.
9. M. P. Davey, T. Maddess, and M. V. Srinivasan. The spatiotemporal properties of the Craik-O'Brien-Cornsweet effect are consistent with 'filling-in'. *Vision Res.*, 38:2037–2046, 1998.
10. R. L. Elder and S. Zucker. Evidence for boundary specific grouping. *Vision Res.*, 38:143–152, 1998.
11. C. Enroth-Cugell and J. G. Robson. Functional characteristics and diversity of cat retinal ganglion cells. *Invest. Ophthalmol. Visual Sci.*, 25:250–267, 1984.
12. H. J. M. Gerrits and A. J. H. Vendrik. Simultaneous contrast, filling-in process and information processing in man's visual system. *Exp. Brain Res.*, 11:411–430, 1970.
13. C. D. Gilbert. Circuitry, architecture, and functional dynamics of visual cortex. *Cereb. Cortex*, 3(5):373–386, 1993.
14. A. Gove, S. Grossberg, and E. Mingolla. Brightness perception, illusory contours and corticogeniculate feedback. *Visual Neurosci.*, 12:1027–1052, 1995.
15. S. Grossberg. 3-D vision and figure-ground separation by visual cortex. *Percept. Psychophys.*, 55(1):48–121, 1994.
16. S. Grossberg and N. McLoughlin. Cortical dynamics of three-dimensional surface perception: Binocular and half-occluded scenic images. *Neural Networks*, 10:1583–1605, 1997.
17. S. Grossberg and E. Mingolla. Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentation. *Percept. Psychophys.*, 38:141–171, 1985.
18. S. Grossberg and D. Todorović. Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Percept. Psychophys.*, 43:241–277, 1988.
19. T. Hansen, G. Baratoff, and H. Neumann. A simple cell model with dominating opponent inhibition for robust contrast detection. *Kognitionswissenschaft*, 9(2):93–100, 2000.
20. T. Hansen and H. Neumann. A model of V1 visual contrast processing utilizing long-range connections and recurrent interactions. In *Proc. ICANN*, pages 61–66, Edinburgh, UK, Sept. 7–10 1999.
21. T. Hansen, W. Sepp, and H. Neumann. Recurrent long-range interactions in early vision. In S. Wermter, J. Austin, and D. Willshaw, editors, *Emergent Neural Computational Architectures based on Neuroscience*, LNCS/LNAI. Springer, Heidelberg, 2000. In press.
22. D. Heeger. Normalization of cell responses in cat striate cortex. *Visual Neurosci.*, 9:181–197, 1992.
23. J. A. Hirsch and C. D. Gilbert. Synaptic physiology of horizontal connections in the cat's visual cortex. *J. Neurosci.*, 11(6):1800–1809, 1991.

24. J. M. Hupé, A. C. James, B. R. Payne, S. G. Lomber, P. Girard, and J. Bullier. Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature*, 394:784–787, 1998.
25. H. Komatsu, I. Murakami, and M. Kinoshita. Surface representation in the visual system. *Brain. Res. Cogn. Brain. Res.*, 5(1):97–104, 1996.
26. D. Mumford. Neural architectures for pattern-theoretic problems. In C. Koch and J. L. Davis, editors, *Large-scale neuronal theories of the brain*. MIT Press, Cambridge, MA,, 1994.
27. H. Neumann. Completion phenomena in vision: A computational approach. In L. Pessoa and P. de Weerd, editors, *Filling-in: From perceptual completion to skill learning*. Oxford Univ. Press. In preparation.
28. H. Neumann. Mechanisms of neural architecture for visual contrast and brightness perception. *Neural Networks*, 9(6):921–936, 1996.
29. H. Neumann, L. Pessoa, and T. Hansen. Interaction of ON and OFF pathways for visual contrast measurement. *Biol. Cybern.*, 81:515–532, 1999.
30. H. Neumann, L. Pessoa, and T. Hansen. Visual filling-in for computing perceptual surface properties. 2000. Submitted.
31. H. Neumann and W. Sepp. Recurrent V1–V2 interaction in early visual boundary processing. *Biol. Cybern.*, 81:425–444, 1999.
32. M. A. Paradiso and K. Nakayama. Brightness perception and filling-in. *Vision Res.*, 31:1221–1236, 1991.
33. L. Pessoa, E. Mingolla, and H. Neumann. A contrast- and luminance-driven multiscale network model of brightness perception. *Vision Res.*, 35(15):2201–2223, 1995.
34. L. Pessoa and H. Neumann. Why does the brain fill in? *Trends Cogn. Sci.*, 2(11):422–424, 1998.
35. L. Pessoa, E. Thompson, and A. Noë. Finding out about filling-in: A guide to perceptual completion for visual science and the philosophy of perception. *Behav. Brain. Sci.*, 21(6):723–802, 1998.
36. T. Poggio, V. Torre, and C. Koch. Computational vision and regularization theory. *Nature*, 317(26):314–319, 1985.
37. D. A. Pollen and S. F. Ronner. Visual cortical neurons as localized spatial frequency filters. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13(5):907–916, 1983.
38. D. C. Rogers-Ramachandran and V. S. Ramachandran. Psychophysical evidence for boundary and surface systems in human vision. *Vision Res.*, 38:71–77, 1998.
39. K. Schmidt, R. Goebel, S. Löwel, and W. Singer. The perceptual grouping criterion of colinearity is reflected by anisotropies of connections in the primary visual cortex. *Europ. J. Neurosci.*, 9:1083–1089, 1997.
40. W. Sepp and H. Neumann. A multi-resolution filling-in model for brightness perception. In *Proc. ICANN*, Edinburgh, UK, Sept. 7–10 1999.
41. A. N. Tikhonov and V. Y. Arsenin. *Solutions of ill-posed problems*. V. H. Winston & Sons, Washington D. C., 1977.
42. R. von der Heydt, E. Peterhans, and G. Baumgartner. Illusory contours and cortical neuron responses. *Science*, 224:1260—1262, 1984.