

A computational model to link psychophysics and cortical cell activation patterns in human texture processing

A. Thielscher · H. Neumann

Received: 13 May 2005 / Revised: 28 August 2006 / Accepted: 6 October 2006 / Published online: 14 November 2006
© Springer Science + Business Media, LLC 2007

Abstract The human visual system uses texture information to automatically, or pre-attentively, segregate parts of the visual scene. We investigate the neural substrate underlying human texture processing using a computational model that consists of a hierarchy of bi-directionally linked model areas. The model builds upon two key hypotheses, namely that (i) texture segregation is based on boundary detection—rather than clustering of homogeneous items—and (ii) texture boundaries are detected mainly on the basis of a large scenic context that is analyzed by higher cortical areas within the ventral visual pathway, such as area V4. Here, we focus on the interpretation of key results from psychophysical studies on human texture segmentation. In psychophysical studies, texture patterns were varied along several feature dimensions to systematically characterize human performance. We use simulations to demonstrate that the activation patterns of our model directly correlate with the psychophysical results. This allows us to identify the putative neural mechanisms and cortical key areas which underlie human behavior. In particular, we investigate (i) the effects of varying texture

density on target saliency, and the impact of (ii) element alignment and (iii) orientation noise on the detectability of a pop-out bar. As a result, we demonstrate that the dependency of target saliency on texture density is linked to a putative receptive field organization of orientation-selective neurons in V4. The effect of texture element alignment is related to grouping mechanisms in early visual areas. Finally, the modulation of cell activity by feedback activation from higher model areas, interacting with mechanisms of intra-areal center-surround competition, is shown to result in the specific suppression of noise-related cell activities and to improve the overall model capabilities in texture segmentation. In particular, feedback interaction is crucial to raise the model performance to the level of human observers.

Keywords Texture processing · Psychophysics · Visual search · Neural model · Context modulation

1 Introduction

Robust segmentation and grouping of the initial visual stimulus into distinct items is crucial for the reliable detection and identification of objects, in particular when being embedded in complex and cluttered environments. It has been proposed that the visual processes underlying segmentation start from the identification of salient discontinuities representing borders between adjacent object surfaces, rather than from the immediate detection of the surfaces *per-se* (e.g., Grossberg, 1980; Grossberg and Mingolla, 1985). For example, in retinal stabilization studies an object seems to disappear and is no longer perceived when its retinal position is held fixed (Krauskopf, 1963; Yarbus, 1967; Safran and Landis, 1998; Martinez-Conde et al., 2004). Instead, the human visual system substitutes the object color by the background color,

Action Editor: Peter Latham

A. Thielscher
Department of Psychiatry, University of Ulm,
Germany

H. Neumann
Department of Neural Information Processing, University of Ulm,
Germany
e-mail: heiko.neumann@uni-ulm.de

A. Thielscher (✉)
Department Ugurbil, Max Planck Institute
for Biological Cybernetics,
Spemannstraße 41,
D-72076 Tübingen, Germany
e-mail: axel.thielscher@tuebingen.mpg.de

indicating that the representation of the object borders gets lost due to neural adaptation processes. The physiological basis which underlies the detection of luminance contrasts signaling putative object borders is well understood. Hubel and Wiesel (1959) demonstrated that simple and complex cells in the primary visual cortex signal the position and orientation of bars (i.e. luminance gradients) in the stimulus and more recent research revealed the existence of contour neurons in V2, capable of completing fragmented contour elements and responding to illusory contours (v.d. Heydt et al., 1984, 1993; Peterhans, 1997). However, the visual system does not utilize luminance contrast alone, but a variety of features for segmentation, such as color, motion, stereoscopic depth and texture. Here, we focus on texture segregation which is demanding as, e.g. natural textures hardly ever stay constant across a surface. Consequently, a comparison of texture elements over wide areas of the visual field is necessary to robustly determine surfaces and surface boundaries. The precise cortical mechanisms underlying texture segmentation still remain largely unknown and, compared to the processing of luminance contrasts, more complex neural mechanisms need to get involved.

A significant amount of our knowledge on human texture segmentation stems from psychophysical studies. In particular, Nothdurft (1991, 1992, 2000c) conducted a comprehensive series of studies in which he characterized (among others) the effects of varying texture density, of element alignment as well as of increasing orientation noise on the detectability of pop-out lines and regions. Tracing back these psychophysical results to their neural origins is a challenging problem. To our knowledge, up to now only Li's modeling studies (2000) have given a putative explanation of some of the observed effects. Her model focuses exclusively on V1 processing and utilizes structural patterns of locally oriented and anisotropic connectivity patterns of excitatory and inhibitory cell interaction. Here, we use a previously developed computational model (Thielscher and Neumann, 2003) to directly link the human behavioral curves characterized by Nothdurft with the model activation patterns. Our model represents a common theoretical framework for a variety of experimental findings on texture processing obtained by electrophysiological, functional imaging and psychophysical methods. The previous study introduced the general model architecture and discussed in detail the specific functionalities of the single components of the model (Thielscher and Neumann, 2003). Now, we focus on key results from Nothdurft's studies and demonstrate through simulations that our model can account for his findings not only qualitatively, but in a quantitative way. Unlike the previous study (Thielscher and Neumann, 2003) in which the model behavior was demonstrated using a few exemplary stimuli (e.g. perfectly homogeneous vs. highly noisy textures), it is shown here that smooth and continuous variations of properties like

texture noise, density or element alignment result in model cell activations which directly correlate with the human psychophysical curves. This enables us to identify the cortical areas and key processing principles which underlie the human behavioral curves in texture segmentation as observed by Nothdurft.

There is now increasing evidence from psychophysical and physiological studies that texture based segmentation starts from the detection of *boundaries* between otherwise homogeneous regions, comparable to the segmentation based on luminance contrasts. For example, single- and multi-unit recordings in monkey primary visual cortex reveal an enhanced activity of orientation selective cells when texture borders defined by an orientation contrast between neighboring line elements fall on their receptive fields (Gallant et al., 1995; Lamme et al., 1998, 1999; Nothdurft et al., 1999, 2000). Furthermore, psychophysical studies demonstrate that texture segmentation depends on the detection of local first-order feature dissimilarities at region borders (Nothdurft, 1985; Landy and Bergen, 1991). Nothdurft (1985, 1991) used patterns composed of oriented texture items to study segmentation based on feature differences in detail. In the stimuli he used, a central pop-out region is embedded in a background composed of texture line elements forming continuous flow patterns (see Fig. 3(A)). Pre-attentive segmentation based on the differences between the line elements of the fore- and background occurs although both regions may contain lines of the same orientation or statistics of oriented items. Critically, successful segmentation depends on the orientation contrast between the texture elements along the border, i.e. on the slope in the gradient of texture orientation. These findings argue in favor of texture processing being based on boundary-finding mechanisms and not on grouping mechanisms which would define surfaces using some criterion for element similarity.

Physiological and neuroimaging studies show that higher visual areas such as V4 play a key role in texture segmentation. For example, lesion studies in monkeys demonstrated that the detection of illusory contours and texture borders as well as the discrimination of texture-defined shapes is severely impaired after ablation of V4 (de Weerd et al., 1996; Merigan, 1996, 2000). Furthermore, using fMRI, it was shown that areas V4 and TEO are significantly stronger activated by texture borders than lower visual areas (Kastner et al., 2000). Higher visual areas are able to integrate information over extended areas of the retinal image (Smith et al., 2001). This capability is crucial for the segmentation of texture stimuli embedded in background orientation noise in order to selectively identify meaningful orientation contrasts which signal the stimuli borders.

In Thielscher and Neumann (2003) we proposed a neural model of pre-attentive segmentation of oriented textures that builds upon the key properties of human texture

processing highlighted above: (i) Texture segmentation is based on boundary detection, and (ii) texture border detection is mainly achieved by the interaction of cells in higher model areas. The model implements texture boundary processing as a distributed task within the hierarchy of ventral visual areas V1, V2 and V4. The model mechanisms were motivated by models of V1–V2 interaction for contour completion and illusory contour formation (Neumann and Sepp, 1999; Grossberg and Raizada, 2000; Ross et al., 2000). The computational capabilities of these approaches are incorporated and preserved by our model, thereby integrating the processing of luminance and texture information for surface boundary finding into a common computational framework.

We focus on the psychophysical studies of Nothdurft (1991, 1992, 2000c) to investigate human performance in the segmentation of oriented textures quantitatively. Nothdurft examined the saliency and detectability of embedded target lines and regions while systematically varying (i) the density and (ii) the alignment of the elements, and (iii) the amount of orientation noise in the stimulus. In this simulation study, we replicate these stimulus variations to directly link the model behavior with the psychophysical curves. This, in turn, allows to identify the putative cortical origins underlying human behavioral performance and highlights the key areas and mechanisms involved. In particular, we demonstrate the importance of recurrent feedback interaction in combination with center-surround competition to enable the model to successfully generate texture boundaries even in presence of significant orientation noise, resulting in a model performance comparable to that of human observers. Furthermore, the psychophysical results of the effects of texture density on the saliency of a target line are traced back to a putative receptive field organization of orientation-selective neurons in V4 as well as their mutual interactions. Finally, effects of texture element alignment on the detectability of a pop-out bar are directly linked to grouping mechanisms incorporated in early visual areas.

The paper is organized as follows. In the next chapter our neural computational model is introduced together with its components and dynamics. Furthermore, the general model behavior is demonstrated considering the model response patterns to two selected stimuli. In chapter 3 the effects of varying texture density on model activation patterns are systematically investigated and a brief discussion outlines the relation of the obtained simulation results to psychophysics. Chapters 4 and 5 demonstrate the effects of texture element alignment and orientation noise, respectively, on the model responses to a pop-out bar, each chapter concluding with a short discussion to demonstrate the link between our results and those of the corresponding psychophysical studies. Finally, a general discussion (chapter 6) summarizes the key results and model features and links the results of a variety of psychophysical, physiological and neuroimaging studies

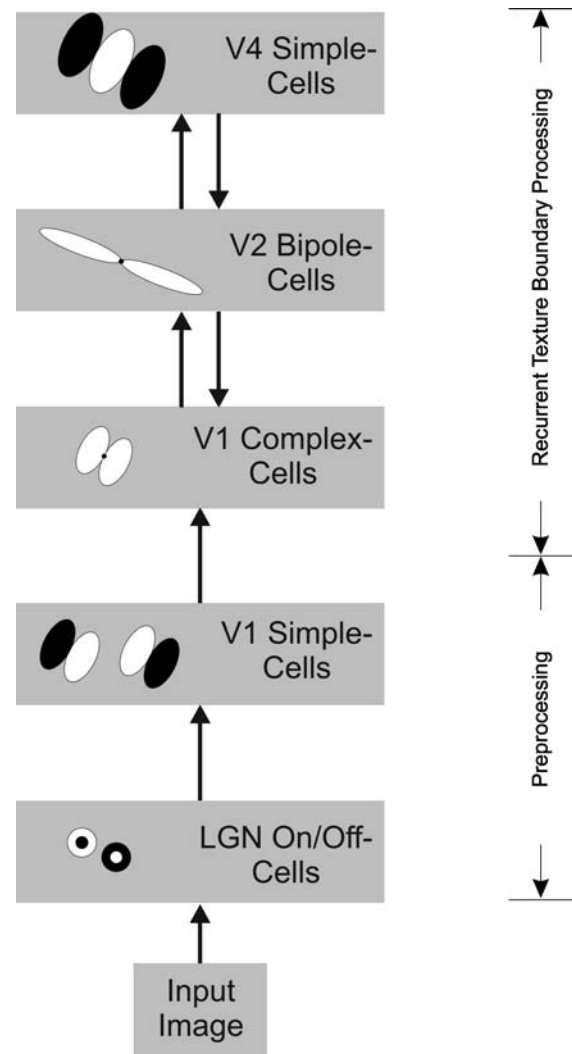


Fig. 1 Outline of the hierarchy of model areas and model cells. Receptive field kernels are sketched in the left column (white: excitatory subfields; black: inhibitory subfields). The field sizes are not drawn to scale

to our model architecture, allowing to integrate them into a common theoretical framework. We conclude with a brief summary.

2 Model description

2.1 General model architecture

A model of recurrent texture boundary processing in the ventral visual stream of static form processing is outlined. According to the known anatomical data, it consists of a hierarchy of bi-directionally linked model areas V1, V2 and V4 (Felleman and van Essen, 1991; Ungerleider and Haxby, 1994; Fig. 1). Each area contains a retinotopically organized map (Serenio et al., 1995) of model cells with gradual

activation dynamics representing the average response (firing-rate) of groups of cortical neurons having similar selectivities. The model cells are orientation selective (Zilles and Clarke, 1997) and their receptive field sizes increase monotonically within the hierarchy of model areas (Smith et al., 2001). Their relative sizes have been parameterized in a ratio $V1:V2:V4 = 1:2.5:8$. The response selectivity of a model cell is determined by the spatial layout and organization of its receptive field pooling the bottom-up cell activities of the previous model area. Model cells belonging to a particular area exhibit stereotyped receptive field organizations and response selectivities. However, receptive field organization differs between the model areas, thereby determining the specific functional role of an area in texture boundary processing.

Feedback connections between model areas enable a recurrent flow of activity and result in a dynamic development of the overall model activation pattern in response to a static stimulus. At each point in time, the activity level of a model cell is determined in three successive computational stages (Fig. 2): First, the bottom-up activity pooled within the receptive field of a cell determines its initial activation. Second, this initial activation level is modulated by excitatory feedback from a model area higher up in the anatomical hierarchy. Feedback is specific to topographical position and orientation, enhancing those cell activities in the lower model area which fit to the activation pattern in the higher area. At the third computational stage, the top-down modulated cell activities undergo intra-areal center-surround competition for contrast enhancement and normalization of activation levels. After onset of stimulus presentation, the three-stage cell dynamics is used to successively update the cell activation levels in all model areas until the overall model reaches its stable final activation pattern.

In the following, the model areas and receptive field organizations are briefly described. Subsequently, the model equations that describe the three-stage cell dynamics are outlined and the general model behavior is demonstrated by probing the network with two selected texture stimuli.

2.2 Model areas and receptive field organization

The model consists of a feed-forward preprocessing stage followed by the recurrent stage of texture border processing (Fig. 1). The preprocessing stage is constituted by model LGN and V1 simple cells. Results of sequential LGN and V1 simple cell processing are fed forward to generate V1 complex cell activity by combining simple cells with opposite polarity selectivity. Complex cells signal the spatial position and local orientation of luminance discontinuities in the input image. This activity is passed on to the recurrent stage of texture border processing which is implemented by bi-directionally linked model areas V1, V2 and V4.

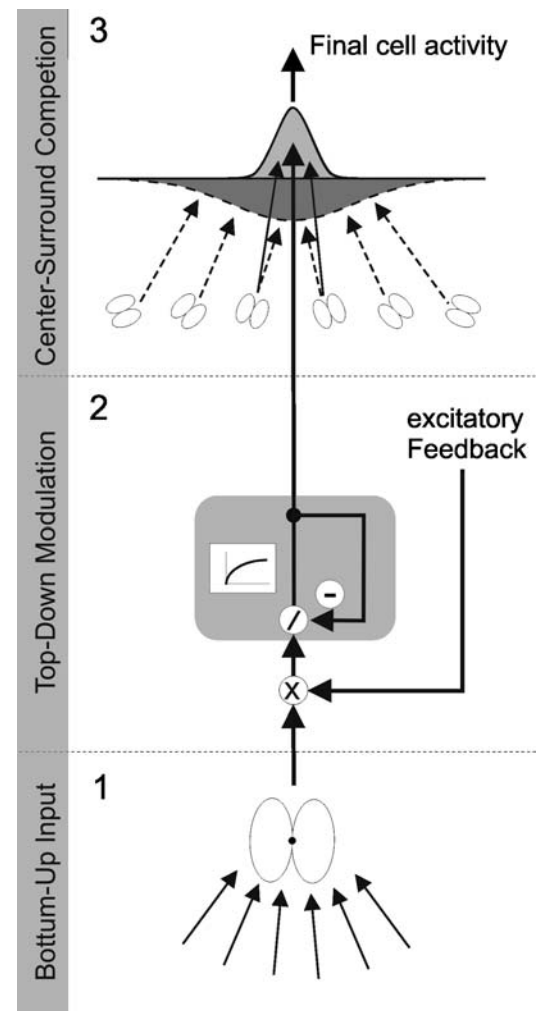


Fig. 2 Three-stage dynamics used to determine the activation level of a cell. (1) The bottom-up input is pooled by the cell's receptive field to generate its initial activation. (2) Multiplicative feedback interaction from higher model areas modulates the initial activation. (3) The top-down modulated activity finally undergoes a stage of shunting ON-center/OFF-surround competition in a spatial and orientational neighborhood

- *Model LGN cells and model V1 simple cells* constitute the first two stages used to initially filter the luminance distribution of the input image in a feed-forward process. Model LGN ON and OFF cells detect local luminance transitions using concentric center-surround receptive fields, which are modeled by a subtractive and half-wave rectified interaction between Gaussian weighted input intensities:

$$\begin{aligned}
 x &= I * (\Lambda_{Center} - \Lambda_{Surround}) \\
 x^{on} &= [x]^+ \\
 x^{off} &= [-x]^+
 \end{aligned} \quad (1)$$

The symbols $\Lambda_{Center/Surround}$ denote isotropic 2D Gaussian kernels (standard deviations $\sigma_{Center} = 0.8$ and $\sigma_{Surround} = 3\sigma_{Center}$ define their sizes in pixels) and $[x]^+$

and $*$ denote half-wave rectification and the convolution operator, respectively. Model V1 simple cells exist for two polarities (dark-light, dl; light-dark, ld) and eight orientations. They have elongated ON and OFF subfields to pool the input delivered by appropriately aligned LGN ON or OFF cells:

$$\begin{aligned} p_{i\theta}^{on/off-left} &= x^{on/off} * \Lambda_{\sigma_x, \sigma_y, 0, -\tau_y, \theta} \\ p_{i\theta}^{on/off-right} &= x^{on/off} * \Lambda_{\sigma_x, \sigma_y, 0, \tau_y, \theta} \end{aligned} \tag{2}$$

The subfields are modeled by anisotropic 2D Gaussian weighting functions $\Lambda_{\sigma_x, \sigma_y, 0, +/-\tau_y, \theta}$. Their size in y - and x -direction is defined by the standard-deviations $\sigma_y = 0.8$ and $\sigma_x = 3.5\sigma_y$. They are shifted perpendicular to their main axis by $+/-\tau_y$ ($\tau_y = 0.8\sigma_y$) and then rotated by θ ($n_{orient} = 8$ different orientations are used with $\theta = 0, \pi/n_{orient} \dots, (n_{orient} - 1)\pi/n_{orient}$). The activity of the subfields is subsequently fed into a soft-AND-circuit which combines additive and multiplicative interactions between the subfield activities to determine the response s of the simple cell and results in a more localized simple cell activation at luminance borders compared to the pure summation of subfield responses (please refer to Neumann et al., 1999 for details). For example, the activation $s_{i\theta}^{ld}$ of a V1 simple cell sensitive for light-dark polarity is given by:

$$s_{i\theta}^{ld} = \frac{A_s(p_{i\theta}^{on-left} + p_{i\theta}^{off-right}) + 2B_s p_{i\theta}^{on-left} p_{i\theta}^{off-right}}{A_s D_s + E_s(p_{i\theta}^{on-left} + p_{i\theta}^{off-right})} \tag{3}$$

Spatial locations in the topographical maps are expressed by the index i . The first term in the numerator denotes the additive excitatory interaction between the ON and OFF subfields, while the second term defines the amount of activation resulting from their multiplicative interaction (parameters $A_s = 1.0$ and $B_s = 10000.0$ control the impact of the additive and multiplicative terms). The denominator normalizes the response by the total sum of subfield responses (the relative strength of the divisive component is controlled by $D_s = 0.05$ and $E_s = 100.0$). Taken together, the first two model stages emulate roughly the functionality seen in the parvocellular layer of LGN and simple cells in V1 (Hubel and Wiesel, 1959, 1962).

- **Model V1 complex cells** form the lowest level of recurrent texture border processing in our model. Model complex cells pool the activity of two simple cells of opposite polarity at each position using half-wave rectified differences between simple cells of opposite polarities:

$$\begin{aligned} c_{i\theta}^{V1} &= A_c([s_{i\theta}^{ld} - s_{i\theta}^{dl}] + [s_{i\theta}^{dl} - s_{i\theta}^{ld}]^+) \\ A_c &= 0.1 \end{aligned} \tag{4}$$

In combination, the hierarchical computation performed in sequence by model LGN and simple cells finally result in complex cell activity which is sensitive to orientation but insensitive to the direction of contrast. The output of the model V1 complex cells thus resembles that of real cortical complex cells (Hubel and Wiesel, 1962). This output activation is subsequently modulated by excitatory top-down feedback from model area V2 and intra-areal center-surround competition, according to the three-stage model cell dynamics motivated above (please refer also to chapter 2.3). The resulting output activation of the V1 complex cells is denoted by $l_{V1}^{(2)}$.

- **Model V2 bipole cells** use two prolated subfields $K^{left/right}$ aligned along the axis of the cell's orientation preference to pool the input delivered by appropriately aligned V1 complex cells. V1 complex cell activity $l_{V1}^{(2)}$ (represented by a 3D matrix: 2D—space, 1D—orientation) is blurred in the orientation domain (through a convolution using a 1D isotropic Gaussian kernel ψ_f with $\sigma_{f_orient} = 0.3$) and convolved with the 2D spatial weighting functions $K^{left/right}$ to determine the subfield activities $f^{left/right}$:

$$\begin{aligned} f^{left} &= l_{V1}^{(2)} * \Psi_f * K^{left} \\ f^{right} &= l_{V1}^{(2)} * \Psi_f * K^{right} \end{aligned} \tag{5}$$

In other words, the convolved orientational and spatial weighting functions $\Psi_f * K^{left/right}$ define two 3D kernels which determine the bottom-up connectivity pattern from the V1 complex cells to the V2 bipole cells. The subfield activities $f^{left/right}$ are determined by convolution of V1 complex cell activity $l_{V1}^{(2)}$ with these 3D kernels. The weighting functions $K^{left/right}$ determining the spatial layout of the subfields are modeled as anisotropic Gaussians, which are cut off in the central part of the cell by means of a sigmoid function (for details please refer to Appendix A). The subfield activations are combined using a soft-AND-gate, in turn only generating significant initial V2 cell responses when both subfields are excited simultaneously:

$$c_{i\theta}^{V2} = \frac{A_t(f_{i\theta}^{left} + f_{i\theta}^{right}) + 2B_t f_{i\theta}^{left} f_{i\theta}^{right}}{A_t D_t + E_t(f_{i\theta}^{left} + f_{i\theta}^{right})} \tag{6}$$

The parameters are $A_t = 2.3$; $B_t = 2600.0$; $D_t = 0.15$; $E_t = 100.0$. In accordance to the three-stage model cell dynamics, the initial V2 cell activities subsequently undergo feedback modulation by V4 activity and intra-areal center-surround competition. As the V2 bipole cells pool the activity delivered by several appropriately aligned V1 complex cells, they are also responsive to luminance contrasts, but have larger receptive fields. Furthermore, their elongated subfields enable them to bridge missing contour information and to recreate a complete shape outline from

fragmented contour parts. The same mechanisms also results in the induction of illusory contours, thus resembling the functional properties of contour neurons in V2 (v.d. Heydt et al., 1984, 1993; Peterhans, 1997). The processing principles implemented by the model bipole cells are motivated by previous models of recurrent V1–V2 interaction for robust contour formation in early vision (Grossberg and Mingolla, 1985; Neumann and Sepp, 1999; Ross et al., 2000; see Neumann and Mingolla, 2001 for an overview of models). The computational capabilities of these approaches are incorporated and enhanced in our model.

- **Model V4 cells** represent the highest level in our model of texture border processing. The orientation of texture elements changes abruptly at region borders, causing activity transitions at corresponding topographical positions in the 2D maps of orientation selective model V1 and V2 cells. Model V4 cells measure the differences between the V2 bipole cell activities $l_{V2}^{(2)}$ pooled by an excitatory center field q^{center} and left and right inhibitory subfields $q^{left/right}$ (modeled as anisotropic Gaussian kernels). Consequently, they respond to gradients of orientation activity at borders while being insensitive to homogeneous regions of same oriented items. Discontinuities in the activity distribution can occur in each input orientation field (depending on the orientation of the elements that define the texture stimulus) and can themselves have different orientations (depending on the orientation of the texture boundaries). Accordingly, a range of model V4 cells with differently oriented receptive fields was defined for each input orientation: For each of the eight V1 complex and V2 bipole cell orientations (denoted by the angle θ), V4 cells in eight orientations exist (denoted by the angle φ ; $n_{orient} = 8$; $\varphi = 0, \pi/n_{orient}, \dots, (n_{orient} - 1)\pi/n_{orient}$). This results in a matrix of $8 \times 8 = 64$ V4 cells in total. Model V4 cells are selective to the orientation of the line elements and to the orientation of the texture border. Texture border orientation is signaled by the V4 activity distribution along the orientation axis denoted by angle φ .

The initial activation of a V4 cell is given by the differences between the V2 bipole cell activities $l_{V2}^{(2)}$ pooled by the center field and the left and right inhibitory subfields. The sum of the two half-wave rectified differences is the initial activation of the cell:

$$c_{i\theta}^{V4} = [q_{i\theta\varphi}^{center} - Cq_{i\theta\varphi}^{left}]^+ + [q_{i\theta\varphi}^{center} - Cq_{i\theta\varphi}^{right}]^+ \tag{7}$$

with:

$$\begin{aligned} q_{\varphi}^{center} &= l_{V(2)}^{(2)} * \Psi_{\varphi} * \Lambda_{\sigma_{q-x}, \sigma_{q-y}, 0, 0, \varphi} \\ q_{\varphi}^{left} &= l_{V(2)}^{(2)} * \Psi_{\varphi} * \Lambda_{\sigma_{q-x}, \sigma_{q-y}, 0, -\tau_{q-y}, \varphi} \\ q_{\varphi}^{right} &= l_{V(2)}^{(2)} * \Psi_{\varphi} * \Lambda_{\sigma_{q-x}, \sigma_{q-y}, 0, \tau_{q-y}, \varphi} \end{aligned} \tag{8}$$

All subfields are rotated by φ and the lateral inhibitory fields are shifted perpendicularly to their main axis by $+/-\tau_{q-y}$ ($\sigma_{q-x} = 22.0$; $\sigma_{q-y} = 8.0$; $\tau_{q-y} = 16.0$; $C = 1.25$). The bottom-up activity is blurred in the orientation domain (using an isotropic Gaussian ψ_{φ} with $\sigma_{q-orient-\theta} = 0.4$). As V4 constitutes the highest model area, the initial activation $c_{i\theta\varphi}^{V4}$ is not modulated via feedback, but is directly fed into the center-surround competition denoted by Eq. (11). In this competition, all V4 cells sensitive to the same V1 orientation interact in a spatial and orientational (denoted by φ) neighborhood. This results in a normalization of the cell activities and the strengthening of those V4 cells which signal the most likely orientation of the texture border. No interaction between different orientation channels exists in model area V4. The final responses $l_{V(4)}^{(2)}$ of all V4 cells sensitive to the same V1 orientation are summed up after center-surround competition and fed back to V2:

$$l_{V4i\theta}^{(2)} = \sum_{k=1}^{n_{orient}} l_{V4i\theta(k-1)\pi/n_{orient}}^{(2)} \tag{9}$$

2.3 Model cell dynamics

The activation of a model cell in response to an input stimulation is determined in three successive computational stages (Fig. 2), namely (i) the pooling of bottom-up activity followed by (ii) activity modulation via feedback interaction and (iii) intra-areal center-surround competition. The pooling of bottom-up activity is implemented as a feed-forward process that is determined by the receptive field organization of a cell, as described in the previous section. The second and third stages of the cell dynamics are denoted by first-order differential equations, which are outlined in the following. In particular, the functional impact of the feedback connections as well as the physiological and theoretical foundations of the proposed scheme of cell dynamics are discussed. For purposes of comprehensibility, the presentation is restricted to the equilibrated responses ($\frac{\partial}{\partial t}l = 0$).

- The modulatory top-down interaction of the second stage in model areas V1 and V2 is denoted by the equation

$$l_{i\theta}^{(1)} = \frac{\beta_1 c_{i\theta} [1 + Ch_{i\theta}]}{\alpha_1 + \gamma_1 c_{i\theta} [1 + Ch_{i\theta}]} \tag{10}$$

The input $c_{i\theta}$ is given by the bottom-up activity weighted by the receptive field kernel (first processing stage). It is sensitive to spatial location i and to orientation θ . Excitatory feedback activity $h_{i\theta}$ delivered by descending cortical pathways modulates the initial input activation. Feedback is sensitive to spatial location and orientation and its strength is controlled by gain factor C . This modulatory

feedback mechanism generalizes the approach of lateral feature linking proposed by Eckhorn et al. (1990) for the synchronization of distributed neural assemblies in cortex. In our model, feedback activity $h_{i\theta}$ multiplicatively enhances the initial activation $c_{i\theta}$ and is only effective at positions with non-zero initial activation. This prevents unspecific activity to spread unintentionally within the topographical map. The denominator term $\gamma_1 c_{i\theta} [1 + Ch_{i\theta}]$ is included in order to account for an upper saturation level of activity $l_{i\theta}^{(1)}$, which is given by the ratio β_1/γ_1 . α_1 determines the rate of activity decay. The numerical values of the constants as used in the simulations are given in Table A.1.

- The third computational stage exerts a scheme of shunting ON-center/OFF-surround competition, using the top-down modulated activity $l^{(1)}$ of the second stage as input to determine the final cell activation level $l^{(2)}$:

$$l_{i\theta}^{(2)} = \frac{\beta_2 \{l^{(1)} * \Psi^+ * \Lambda^+\}_{i\theta} - \delta_2 \{l^{(1)} * \Psi^- * \Lambda^-\}_{i\theta}}{\alpha_2 + \zeta_2 \{l^{(1)} * \Psi^- * \Lambda^-\}_{i\theta}} \tag{11}$$

This equation is based on a “Mexican Hat” shape for spatial as well as orientational information. Lateral interaction of neighboring cells is mimicked by convolution of the top-down modulated activity $l^{(1)}$ with the Gaussian weighting functions Λ^+ , ψ^+ , Λ^- and ψ^- . The convolutions of the kernels $\{\psi^+ * \Lambda^+\}$ and $\{\psi^- * \Lambda^-\}$, respectively, determine the excitatory and inhibitory 3D connectivity patterns between neighboring bipole cells. The letters Λ^+ and ψ^+ denote excitatory weighting functions which pool activity of those cells in a close spatial neighborhood (Λ^+) which have a similar orientation preference (ψ^+). The inhibitory kernels Λ^- and ψ^- pool activity in a wider spatial and orientational neighborhood. Convolution of activity $l^{(1)}$ with the Gaussian weighting functions results in two 3D blocks $\{ \}_{i\theta}$ in which the excitatory and inhibitory surround activity for a cell at spatial position i and orientation θ is represented at the corresponding 3D position.

The constant δ_2 controls the strength of subtractive inhibition by surround activity given by the term $\delta_2 \{l^{(1)} * \Psi^- * \Lambda^-\}_{i\theta}$. In addition, shunting interaction is employed to incorporate divisive inhibition by the term $\zeta_2 \{l^{(1)} * \Psi^- * \Lambda^-\}_{i\theta}$. Shunting interaction achieves a contrast-enhancement and normalization of the initial top-down modulated activities, resulting in an activity dependent tuning of the cells’ responsiveness.

In all, the three stages of computation realize a *context-selective gain enhancement* or *soft-gating mechanism*: Initial cell activities, which match the activation pattern in the next model area, are enhanced via excitatory feedback connections and in turn inhibit cells in their neighborhood via

center-surround competition. This results in a facilitation of bottom-up activity matching the “expectation” of the cells in the higher model area (Grossberg, 1980; Mumford, 1994). As cells in higher areas integrate information over wide parts of the input image, the overall process of recurrent activity interaction enables a context-selective enhancement of salient texture arrangements, while at the same time spurious and perceptually irrelevant activities are suppressed. The computational behavior essentially realizes a *biased competition mechanism* such as suggested in the context of attentional selection (Reynolds et al., 2000). Reynolds demonstrated that spatial attention shifts a cell’s overall response profile to increase its sensitivity to stimuli of low saliency. In our model cell dynamics, feedback has the same effect on the cells’ input-output functions. Our recursive scheme of cell interaction draws upon basic ideas of the architecture proposed by Finkel and Edelman (1989) using “reentry” connections. In their model excitatory feedback is utilized to enhance but also to induce new activity in the lower areas. Furthermore, it utilizes a mixture of excitatory and inhibitory feedback connections. In contrast, our model employs a stereotyped pattern of (weak) excitatory feedback modulating the activity in lower areas, followed by shunting inhibition. Using modulatory instead of driving feedback connections enhances the stability of the overall network and prevents “hallucinations” in the model activation patterns. Accordingly, Crick and Koch (1998) suggested in their “no-strong-loops hypothesis” that a directed loop between two cortical areas will *not* consist of two driving connections, but will use one driving and one modulatory connection in order to avoid uncontrolled oscillations of the overall system and to limit the amount of inhibition necessary to achieve a stable network behavior. Indeed, several physiological studies indicate that feedback from higher visual areas is not capable of driving cells in lower areas, but *modulates* their activity (Sandell and Schiller, 1982; Mignard and Malpeli, 1991; Salin and Bullier, 1995; Hupé et al., 1998; Przybyszewski et al., 2000). For example, feedback activation alone is not sufficient to drive V1 neurons if they are not stimulated by a visual feeding input (Sandell and Schiller, 1982). Likewise, projections from the striate cortex to the LGN multiplicatively enhance responses of parvocellular neurons to grating stimuli (Przybyszewski et al., 2000). Hupé et al. (1998) demonstrated that feedback from area V5 (or MT) to areas V1, V2 and V3 amplifies and facilitates neural responses in these areas. In particular, in their study, feedback from V5 had the highest impact on cell firing rates in V3 for low saliency stimuli, but had a markedly weaker effect for high saliency stimuli. Our model dynamics mimics this behavior by the saturation properties of the employed cell dynamics. Given a strong bottom-up input the output of a model cell is driven to almost saturate. Consequently, feedback activity is hardly effective in further modulating the cell activity to get further enhancement. For

weak bottom-up input, on the other hand, feedback activity can substantially amplify the activation strength of the corresponding model cell.

The sequence of activity enhancement via feedback followed by center-surround competition has some advantages compared to a model utilizing stages in reversed order (i.e., first center-surround competition followed by feedback interaction). In a hierarchy of feed-forward connected cortical areas, it is commonly proposed that lateral inhibition is necessary to keep the activity in higher areas focused and narrow despite the larger receptive field sizes in these areas (Kastner et al., 2001). We suggest that the same holds for the integration of feedback activity from higher areas that is delivered to lower areas. Consider, for example, the case that the activity distribution in a higher model area is somewhat “blurred” or broadened due to the larger receptive field sizes of the cells in that area. In a first step, feedback from this area will enhance the cell activity in a lower area at topographical positions corresponding to the same blurred region. Without center-surround competition following the feedback interaction, this in turn will lead to an activity pattern in the higher area, which tends to be even more unfocused. However, that effect is prevented by center-surround competition in the lower area narrowing the zone of activity enhancement via feedback and, in consequence, helping to spatially focus the responses to an input pattern in both the lower and the higher model area.

Prior to the actual simulations, the constants of the equations defining the model cell dynamics were empirically determined in such a way that the whole network could reach a stable activation pattern quickly after onset of input pattern presentation. In order to speed up processing, the differential equations were solved at equilibrium in response to a constant input. Initially, the activities of all model layers were set to zero. The input image was clamped and the activities of the model areas were sequentially updated. Starting with model area V1 the activities of the model stages were successively determined one after the other each time applying the three-stage cell-dynamics outlined above (Fig. 2). This was repeated several times until the overall model reached a final, stable final activation pattern which was typically the case after 4–5 iterative cycles. Each simulation was continued until iteration 9 in order to visually demonstrate the stability of the solution. A comparison with results obtained by numerical integration of the model equations (using Euler’s method) revealed that the use of equilibrium responses did not affect the results of the final activation patterns. The values for the constants and the standard deviations of the Gaussian kernels used in the simulations are given in Table A.1. They were kept constant through all numerical experiments and are identical to those used in Thielscher and Neumann (2003). In order to circumvent numerical boundary artifacts and to guarantee well-defined convolution results at image

boundaries, white pixels were padded to each side of the input image prior to stimulus presentation. The number of rows and columns padded was three times the standard deviation of the widest Gaussian used. This is equivalent to presenting the stimulus on a uniform white background.

2.4 General model behavior

In the following, the general model behavior is exemplarily demonstrated using the final equilibrated model activation patterns in response to two selected texture stimuli, as shown in Fig. 3. For purposes of visualization, the activation patterns were summed up over all orientations θ (and φ in case of area V4) at each spatial position and the resulting two-dimensional activity distributions were illustrated as gray-scale images, with the maximal activity of each model area coded as white and no activity coded as black.

The stimulus of Fig. 3(A) consists of a texture array composed of oriented line items in which a central vertical bar pops out, caused by an orientation contrast (OC) of 50° between the line elements of the bar and the texture background (image size 270×270 pixels). Additionally, a constant orientation shift of 10° between neighboring lines (background noise, BN) leads to apparent “flow patterns” and introduces orientation noise. Model V1 complex cells signal the outlines and orientations of the line elements of the stimulus arrangement. This V1 activity pattern serves as bottom-up input to model V2 bipole cells, which group it to form continuous smooth curves. The resulting V2 activation pattern corresponds well with the subjective impression of apparent “flow patterns” visible in the stimuli. Finally, the borders of the bar and the entire texture field are detected by model V4 cells, which respond to orientation discontinuities in the input delivered by model area V2. In the *fully recurrent model* (Fig. 3(A) bottom row), the responses of model V2 cells at the borders are enhanced via modulatory feedback interaction from V4. This enhancement in turn results in stabilized and increased V4 cell activities signaling the outline of the pop-out bar, as can be seen when comparing the response patterns of the recurrent and the feed-forward model. In the recurrent model, the enhancement of V2 cell activity via feedback from V4 leads to an indirect and weaker modulation of model V1 cells in the range of 10–20% of their mean firing rate (this weak influence is hardly visible in the gray scale pictures; please refer to Thielscher and Neumann, 2003 for quantitative results).

The stimulus of Fig. 3(B) consists of a field of regularly spaced vertical lines, superimposed by two oblique line elements representing an X (image size 270×270 pixels). The model V4 cells signal the borders of the field of vertical lines, but are insensitive to its homogeneous inner region. Additionally, they react to the orientation discontinuity caused by the X pattern. The functional impact of feedback from

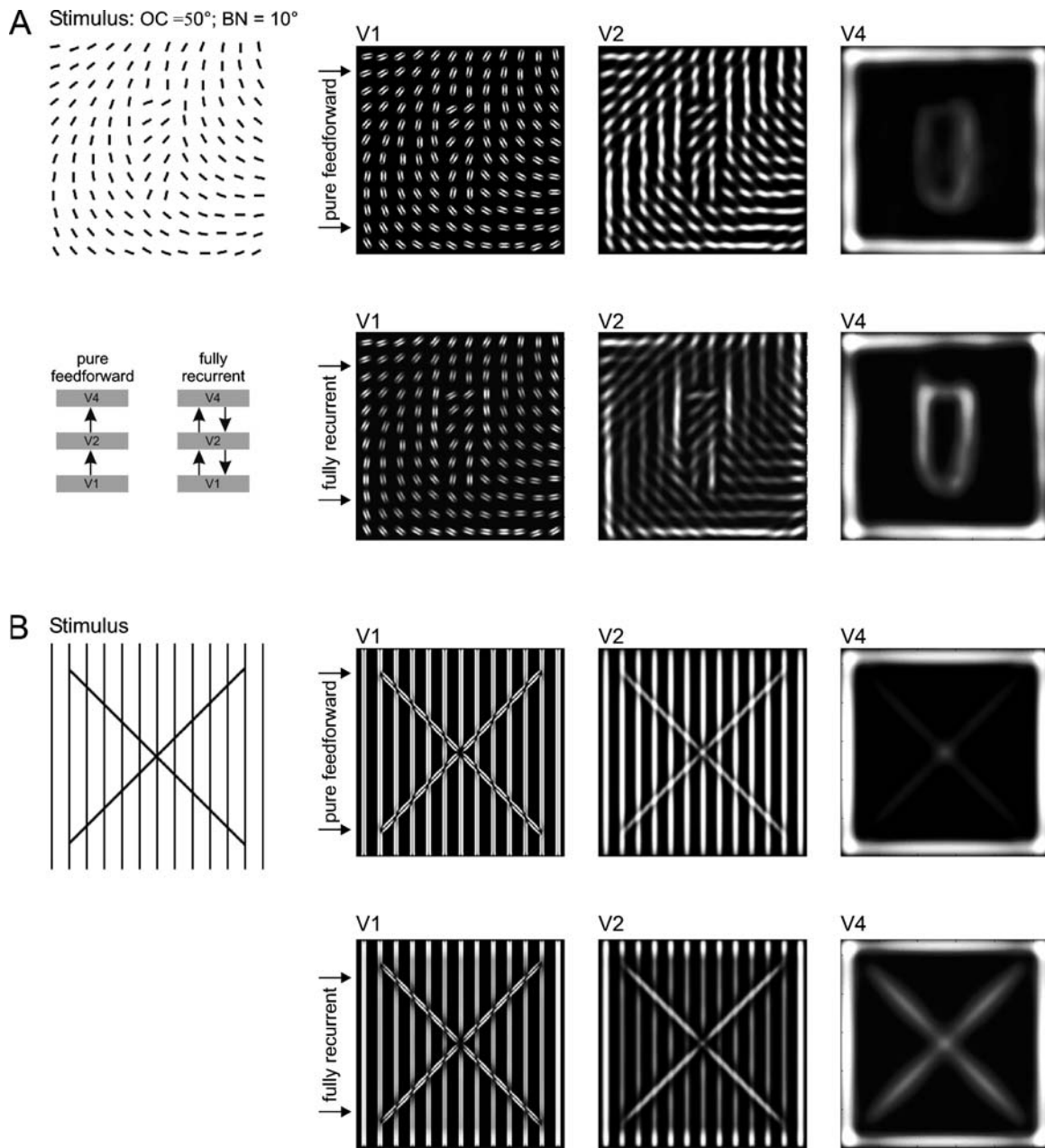


Fig. 3 Equilibrated cell activation patterns in response to two selected stimuli, summed over all orientations for purposes of visualization. (A) The borders of the texture region and the pop-out bar are detected by model V4 cells. In the fully recurrent model (bottom row), feedback from V4 results in enhanced model V2 cell responses at topographical positions of high V4 activity. Compared to the pure feed-forward model (top row), this enhancement helps to stabilize and increase V4 cell activities signaling the outline of the pop-out bar. The maximal activation strength (in arbitrary units) are $\max\{V1, V2, V4\} = \{0.048, 0.029,$

$0.049\}$ for the feedforward and $\{0.122, 0.105, 0.121\}$ for the recurrent model. (B) The borders of the field of vertical lines as well as the two oblique lines resembling an X are detected by model area V4. In the recurrent model (bottom row), feedback interaction helps to segregate the model V2 activities signaling the X from the activities due to the field of vertical lines. This in turn results in increased V4 responses to the X. Maximal activation strength: $\max\{V1, V2, V4\} = \{0.043, 0.053, 0.070\}$ (feedforward) and $\{0.115, 0.118, 0.123\}$ (recurrent)

V4 is revealed by comparison of the activation patterns of the recurrent and feed-forward model variants: Model V2 cell activities which signal the presence of the X pattern are enhanced, helping to segregate them from the V2 activities due to regular vertical lines. The enhanced V2 activities, in

turn, result in markedly increased V4 activities to signal the X pattern.

The model responses to the two selected stimuli exemplarily demonstrate the main impact of feedback connections that enable a recurrent flow of activity: They help to focus the

processing of the overall model on salient discontinuities, thereby enhancing and stabilizing the corresponding V4 activation patterns. A systematic investigation of the impact of feedback on model performance is subject of chapter 5.

3 Variation of texture density

In this chapter, the effect of variation of texture density (as studied by Nothdurft, 2000c) on model activation patterns is determined by comparing the simulation results in response to homogenous texture fields of varying density, with a central target line being either present or absent.

3.1 Stimuli & analysis of the model activation pattern

Gray scale images of sizes 300×300 pixels were used as stimuli containing texture arrays composed of oblique black lines (45°) on a white background. The line length was fixed at 8 pixels and the elements' center-to-center distances were systematically varied using spacings of $\{14, 17, 20, 23, 28, 35, 47, 71\}$ pixels (see Fig. 4(A)). Line positions were slightly varied to avoid alignment effects. Two sets of simulations were performed, one with and one without a central target element defined by an orientation contrast (OC) of 90° . The mean inner V4 activity (Fig. 4(B) and (C): region defined by the dashed white lines) was determined and plotted against element distance for both sets of simulations.

3.2 Results

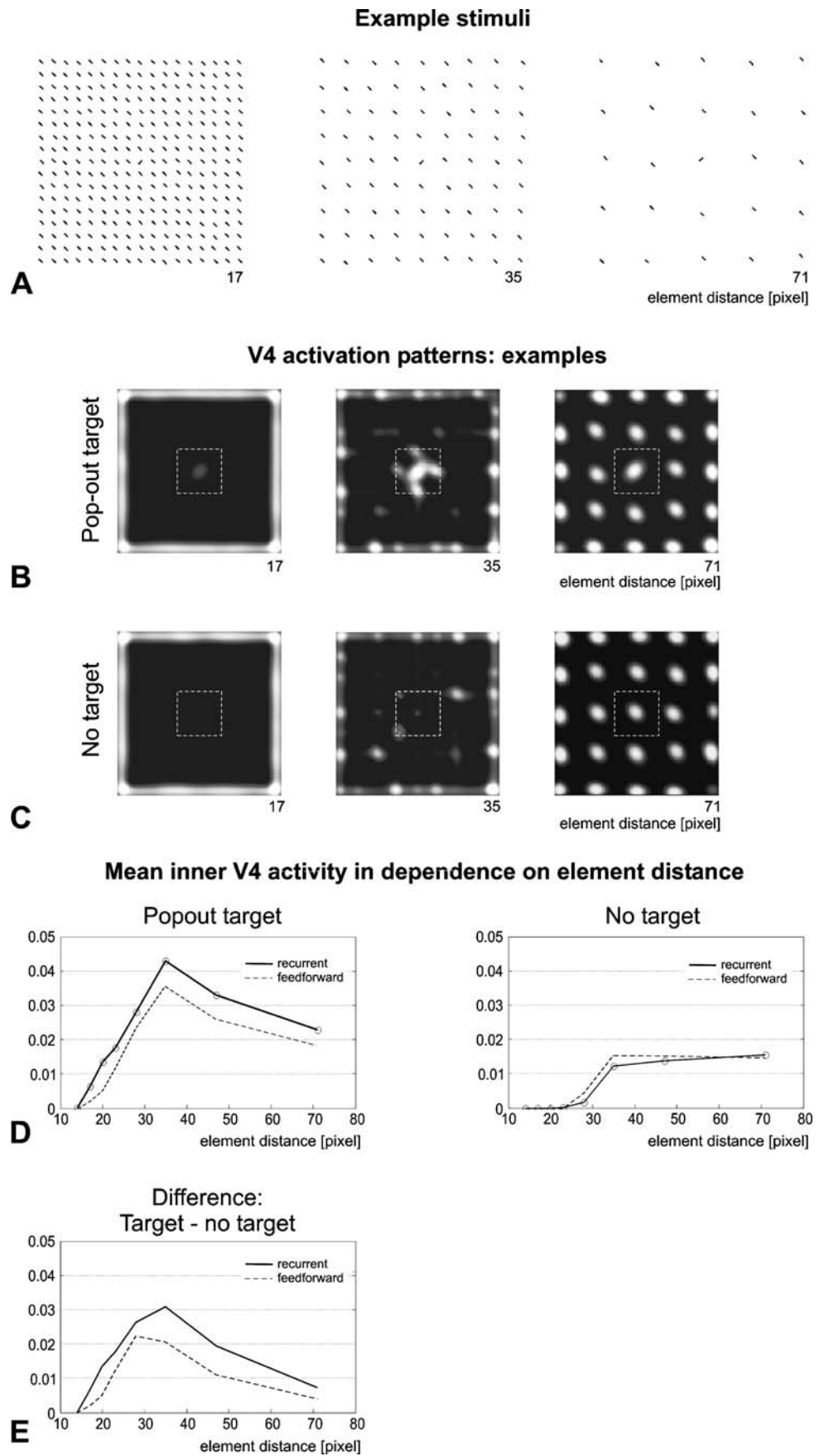
- *Simulations with target element:* For very high texture densities, the intra-areal center-surround competition between cells in the lower model areas results in attenuated V1 and V2 cell responses to all line elements, except to those at the outer texture borders. This effect also diminishes the V1 and V2 responses to the target line element. In consequence, for high texture densities, model area V4 receives only weak bottom-up activity signaling the presence of a target, in turn resulting in a low mean inner V4 activity (see element distance 17 in Fig. 4(B)). The influence of center-surround competition in model areas V1 and V2 diminishes with increasing element distance, leading to a clear detection of the target line by model V4 cells. Furthermore, model V4 cells also respond to the background lines surrounding the target. This effect is caused by a partial loss of lateral inhibition by an iso-oriented line at the position of the target line. The combination of both effects results in a high mean inner V4 activity for medium texture densities (element distance 35 in Fig. 4(B)). For low texture densities, only the V4 activity caused by the target line remains within the inner V4 region defined by the dotted lines, resulting in a decay of mean inner V4 activ-

ity (element distance 71 in Fig. 4(B)). Taken together, the mean inner V4 activity exhibits an approximate inverted U-shaped dependency of response on element distance in presence of a target line (Fig. 4(D) left). In our simulations, the size and shape of the inner V4 region was determined somewhat arbitrarily. Altering the region will rescale the curve and shift the point of the maximum. However, the general shape of the curve will remain inversely U-shaped with its maximum at medium texture densities. For comparison, the results of the feedforward model are depicted in Fig. 4(D) (left). Unlike the stimuli in Fig. 3, the texture density stimuli do not contain orientation noise or crossing lines and, consequently, the impact of a recurrent flow of activity on model V4 activation strength is only modest.

- *Missing target element:* In case of the target line being absent, the mean inner V4 activity monotonically increases with increasing element distance (Fig. 4(C) and (D) right). This effect is caused by the limited spatial extent of the model V4 receptive fields, particularly of the lateral inhibitory subfields that pool the iso-oriented surround. Moreover, the impact of V4 intra-areal center-surround competition decreases with increasing element distance.
- The specific effect of a pop-out target on V4 cell activation is determined by taking the difference between the mean inner V4 activities with and without a target line and plotting the resulting values against element distance (Fig. 4(E)). The resulting curve has a clear peak for medium texture densities.

Fig. 4 Effects of varying texture density on model V4 activation patterns. (A) Example stimuli of varying texture density containing a central pop-out line element. (B) Equilibrated V4 activation patterns in response to the example stimuli. In case of very high texture densities (element distance 17), the central pop-out line evokes low model V4 activity in the region defined by the dashed line. Medium texture densities (element distance 35) result in high V4 activity in this region and low texture densities (element distance 71) lead to slightly lower, but still clear-cut activities. However, for low texture densities, the line elements composing the texture background also evoke V4 activity blobs. Maximal activation strength in V4: $\{0.121, 0.137, 0.160\}$ for line distances $\{12, 35, 71\}$. (C) Equilibrated V4 activation patterns in response to stimuli equivalent to those depicted in (A), but with the central pop-out target line being replaced by an iso-oriented background line. The mean inner V4 activity in the region defined by the dashed line increases monotonically with decreasing texture density. Maximal activation strength in V4: $\{0.120, 0.141, 0.159\}$ for line distances $\{12, 35, 71\}$. (D) Mean inner V4 activity in response to a target line (left) or to an iso-oriented background line (right), plotted in dependence on texture density. For comparison, the mean inner V4 activity of a pure feedforward model is plotted as dashed lines. (E) The difference between the mean inner V4 activities of the target and non-target trials reveals the amount of V4 activity which stems specifically from the orientation contrast of the pop-out line. The recurrent and feedforward models exhibit the same qualitative behavior with the maximal difference between target and no target being higher for the recurrent model

Fig. 4



3.3 Discussion

The effect of variations of texture density on the saliency of a pop-out target has been investigated in a psychophysical study (Nothdurft, 2000c). In that experiment, subjects had to rate the saliency of a target defined by an OC of 90° while systematically varying texture density. The key results were:

- The saliency of the pop-out target was low for high densities, it increased to a maximal value for medium densities and decreased again for low density values. In our model, the mean V4 activity in the topographical region surrounding the target element exhibits the same dependency on texture density (Fig. 4(D) left).
- When no pop-out target was present, the iso-oriented texture elements themselves were rated psychophysically as increasingly salient at low texture densities. Likewise, in our model, the inhibitory effect of the iso-oriented surround on the mean inner V4 activity vanishes for low texture densities (Fig. 4(D) right).
- In order to isolate the saliency caused by the orientation contrast, the difference between the saliency ratings of the target and non-target trials was determined and plotted against texture density. The resulting curve increases from low to medium element distances and decreases afterwards to approach zero at high distance values. In our simulations, the curve depicting the difference between the mean inner model V4 activities in response to the target and non-target stimuli exhibited the same general dependency on texture density (Fig. 4(E)).

Using the results of our simulations, the psychophysical data on the effects of texture density (Nothdurft, 2000c) may be interpreted as follows: Let us assume that the V4 activation pattern is “read out” by a decision stage using spatial receptive fields with some kind of center-surround organization. Then the mean inner V4 activity is a gradual measure, or signature, of the “saliency” of the topographical region corresponding to the pop-out target: The higher the mean activity in this region, the higher the probability that the decision stage rates the underlying input image as containing a “salient” element.

As shown in Fig. 4(D) and (E), the feedforward model exhibits the same qualitative behavior as the recurrent model, with the effect of feedback being limited to an increase of the model activation strength in response to the pop-out target. The impact of feedback is only minor in this task which uses homogeneously oriented background lines and a target line exhibiting a high OC. In contrast, intact feedback connections are crucial in case of stimuli containing orientation noise as will be demonstrated in Section 5.

4 Alignment effects

In the following set of simulations, the alignment of texture line elements relative to the border of a pop-out bar was systematically varied in order to quantify the resulting effect on model V4 activation patterns.

4.1 Stimuli & analysis of model activation patterns

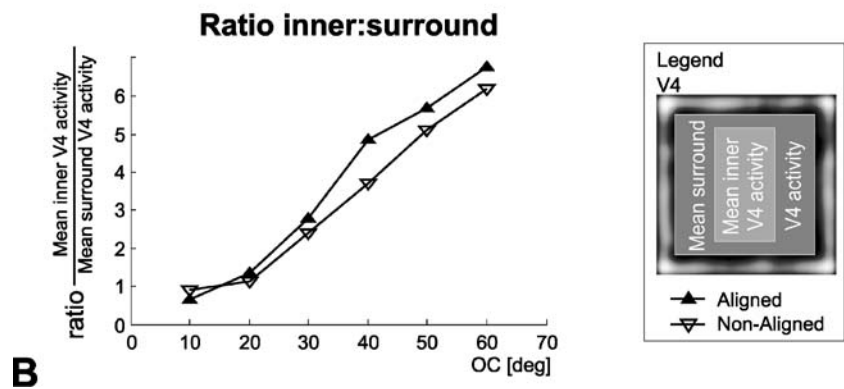
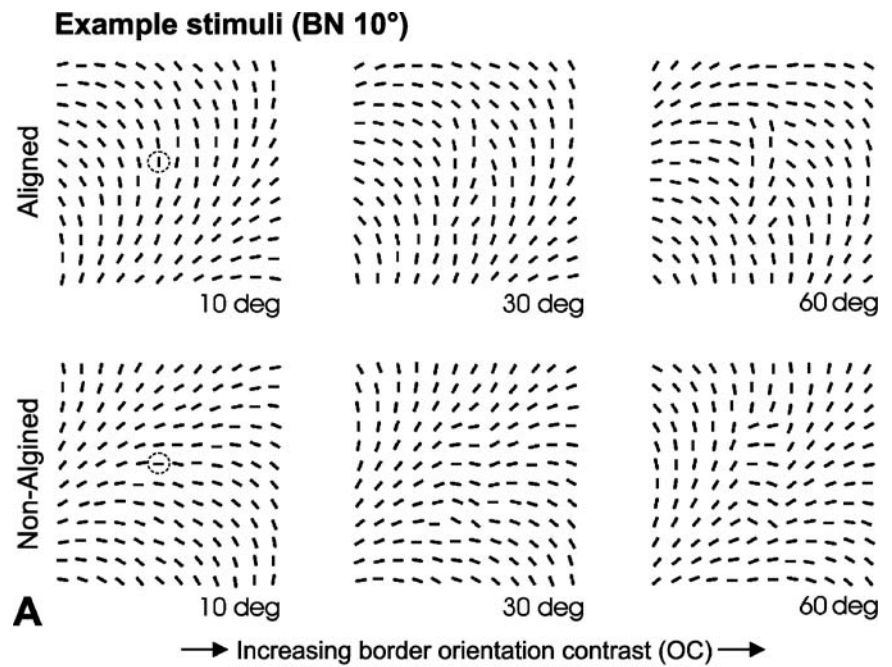
Stimuli were adapted from psychophysical studies (Nothdurft, 1991, 1992) in order to be able to directly compare activation patterns of our model with human performance in texture segmentation. Gray scale images of sizes 270×270 pixels containing texture arrays of 12×12 black line elements on a white background were utilized (line length 12 pixels). In the central region a bar pops out, caused by a certain OC between the neighboring lines at the region border. Line positions were slightly varied to avoid alignment effects. An additional constant orientation shift between neighboring line elements leads to apparent “flow patterns” and thus introduces orientation noise (Fig. 5(A)). The difficulty to detect the central pop-out bar can be controlled by means of the orientation contrast at the border and the background orientation noise (BN), allowing a systematic parametric examination of the activation pattern of the model network.

In the current simulations, quantitative values of mean cell activation levels in model area V4 were obtained. The background orientation noise was fixed at 10° , and the border OC was systematically varied between 10° and 60° in steps of 10° . Two sets of simulations were performed, namely one using texture patterns containing line elements which were maximally aligned with the long border of the pop-out bar, and one with texture arrays consisting of elements which were maximally non-aligned with the border. Alignment was controlled by the orientation of the line element in the gray dotted circle depicted in Fig. 5(A), which was either vertical or horizontal. For each value of OC, the mean cell activation strengths in the model V4 region corresponding to the inner pop-out bar (light gray region of model area V4 depicted in the legend of Fig. 5(B) as well as the region corresponding to the surround (dark gray region) were determined. The ratios between these two values were plotted as a function of OC for both sets of simulations.

4.2 Results

In general, the ratio between the inner and surround V4 activity exhibits a monotonically increasing dependency on OC (Fig. 5(B)). We can assume here that the pop-out bar is detected, when this ratio exceeds a given threshold which is significantly higher than 1 in order to gain noise robustness. For both the aligned and non-aligned condition, this is the

Fig. 5 Effects of element alignment on model V4 activation strength. (A) Example stimuli of varying OC embedded in a BN of 10°. Alignment is controlled by the line element marked by the dashed circle, which is either parallel (aligned) or perpendicular (non-aligned) to the vertical border of the bar. (B) Ratios between mean inner and surround V4 activity in dependence on OC. For OC-values exceeding 20°, the model achieves higher inner-to-surround ratios for aligned stimuli compared to non-aligned stimuli



case for OC-values greater than 30°. However, higher OC-values are necessary for non-aligned stimuli compared to aligned stimuli in order to reach the same inner-to-surround ratio. For example, approximately the same value is achieved for aligned stimuli with an OC of 40° and for non-aligned stimuli with an OC of 50°.

We have demonstrated that long-range groupings mediated by model V2 bipole cells constitute the neural origin of the observed alignment effect (Thielscher and Neumann, 2003). Model V2 cells tend to align the elements of the background and the pop-out bar to continuous activation patterns and thus reduce the initial OC at the border of the bar. As a consequence, the cell activity in the inner V4 region is reduced. In the non-aligned condition, the line elements are oriented approximately perpendicular to the long borders of the bar, enabling the V2 bipole cells to act across these long borders and resulting in an attenuation of inner V4 activity. In contrast, in the aligned condition V2 bipole cells

group line elements only across the short borders of the pop-out bar, which has a markedly weaker effect on model V4 activity.

4.3 Discussion

Nothdurft (1992) investigated the effect of element alignment on the detectability of a pop-out bar. Subjects had to identify pop-out bars embedded either in aligned or in non-aligned stimuli. While keeping the BN fixed at 10°, detectability

- monotonically increased with increasing OC.
- was always higher in the aligned compared to the non-aligned condition for OC-values exceeding 20°.
- differed maximally between aligned and non-aligned stimuli for intermediate OC-values of 30° and 40° (weak visibility of the pop-out bar).

Nothdurft concluded that “the ability of the visual system to detect line continuations may add to border extraction from local contrast” (Nothdurft, 1992, p. 361). In our model, alignment effects lead to spatial long-range groupings by V2 bipole cells which in turn reduce the V4 activity at region borders in the case of a common alignment of the lines across these borders. This results in a ratio between inner and surround V4 activity exhibiting the same dependency on OC and on element alignment as the detectability measure in the psychophysical experiment. However, in our model, the ratio (i.e. the detectability) is decreased in the non-aligned condition rather than increased in the aligned condition. This indicates that the main effect of the ability of the visual system to detect line continuations is more likely to disturb (rather than to support) the border extraction from local contrast. To conclude, the experimental findings of alignment effects that influence the detectability of a central pop-out bar can be traced back to the activity of model V2 bipole cells as the most probable neural origin.

5 Effects of orientation noise

The effect of orientation noise on model performance in texture boundary processing was investigated by systematically varying the orientation contrast at the border of a pop-out bar as well as the overall background orientation noise in the stimulus.

5.1 Stimuli & analysis of model activation patterns

As in the previous experiment, stimuli adopted from psychophysical studies (Nothdurft, 1991, 1992) were used in order to be able to directly compare activation patterns of our model with human performance in texture segmentation. The gray scale images had a size of 250×250 pixels and contained texture arrays of 12×12 line elements. The difficulty to detect the central pop-out bar is controlled by means of the OC at the border and the background orientation noise (BN). The BN was added as additional constant orientation shift between neighboring line elements, leading to apparent “flow patterns”.

In the current set of simulations, BN was systematically varied between 0° and 30° in steps of 5° (example stimuli are depicted in Fig. 6(A)). For each value of BN, input-output curves of the model behavior in dependence on the border orientation contrast were measured. The minimal OC tested was determined by the BN (i.e., in case of $BN = 15^\circ$, the minimal OC was 15°), and the maximal OC was 90° . A stepping of OC of 10° was used, except for BN values of 5° , 15° and 25° , for which the initial stepping was 5° .

As demonstrated in the previous section, model V4 activation strength is influenced by the alignment of the line elements with the contour of the pop-out bar. This effect necessitates a control of the orientation of the line elements in the current set of simulations. For each combination of BN and OC, three different texture patterns with varying orientations of the line elements of the pop-out bar were tested: One texture pattern had elements which were maximally aligned with the long border of the pop-out bar, the other pattern consisted of elements which were maximally non-aligned with the border and the third pattern had elements with an in-between arrangement (see Fig. 5(A) & Nothdurft, 1992, Fig. 2). For further analysis of the model V4 activation pattern, the mean activation level in response to the three stimuli was calculated at each topographical position.

For each combination of BN and OC, the mean cell activation strengths in the V4 region were determined corresponding to the inner pop-out bar (red region of model area V4 depicted in the legend of Fig. 6) as well as the region corresponding to the surround (blue region). These two values were plotted depending on OC for each value of BN. Additionally, the ratio between these two values was calculated and in turn plotted as a function of OC for each value of BN. In order to suppress high ratios resulting from very low mean surround activity levels, a constant α was added (as tonic activation) to the mean activation of the surround region ($\alpha = 0.0154$ and 0.0022 for the recurrent and feed-forward model).

Two model variants were considered, namely a fully recurrent and a pure feed-forward model (see Fig. 3(A)). The comparison of their activation patterns allowed to explore the functional role of the feedback connections in texture boundary processing.

5.2 Results

5.2.1 Recurrent model

The input-output curves of the recurrent model in dependence on OC are depicted in Fig. 6(B)–(E). For a BN of 0° , the mean inner V4 activity exhibits a sigmoidal dependency on OC (Fig. 6(B): purple line with circles) with its steepest increase for OC-values from 10° to 30° . Higher values of BN result in increasingly flatter input-output curves of the inner V4 activity, indicating that the orientation noise tends to disturb the formation of model V4 activity that signals the border of the pop-out bar. In contrast to the sigmoidal curve of $BN = 0^\circ$, the curve of, e.g. the highest BN value of 30° increases approximately linearly from 0.024 to 0.028 (Fig. 6(B): green line with triangles).

In Fig. 6(C) the dependencies of the mean V4 surround activity on OC are depicted. For a specific value of BN, the surround activity remains at a more or less constant level

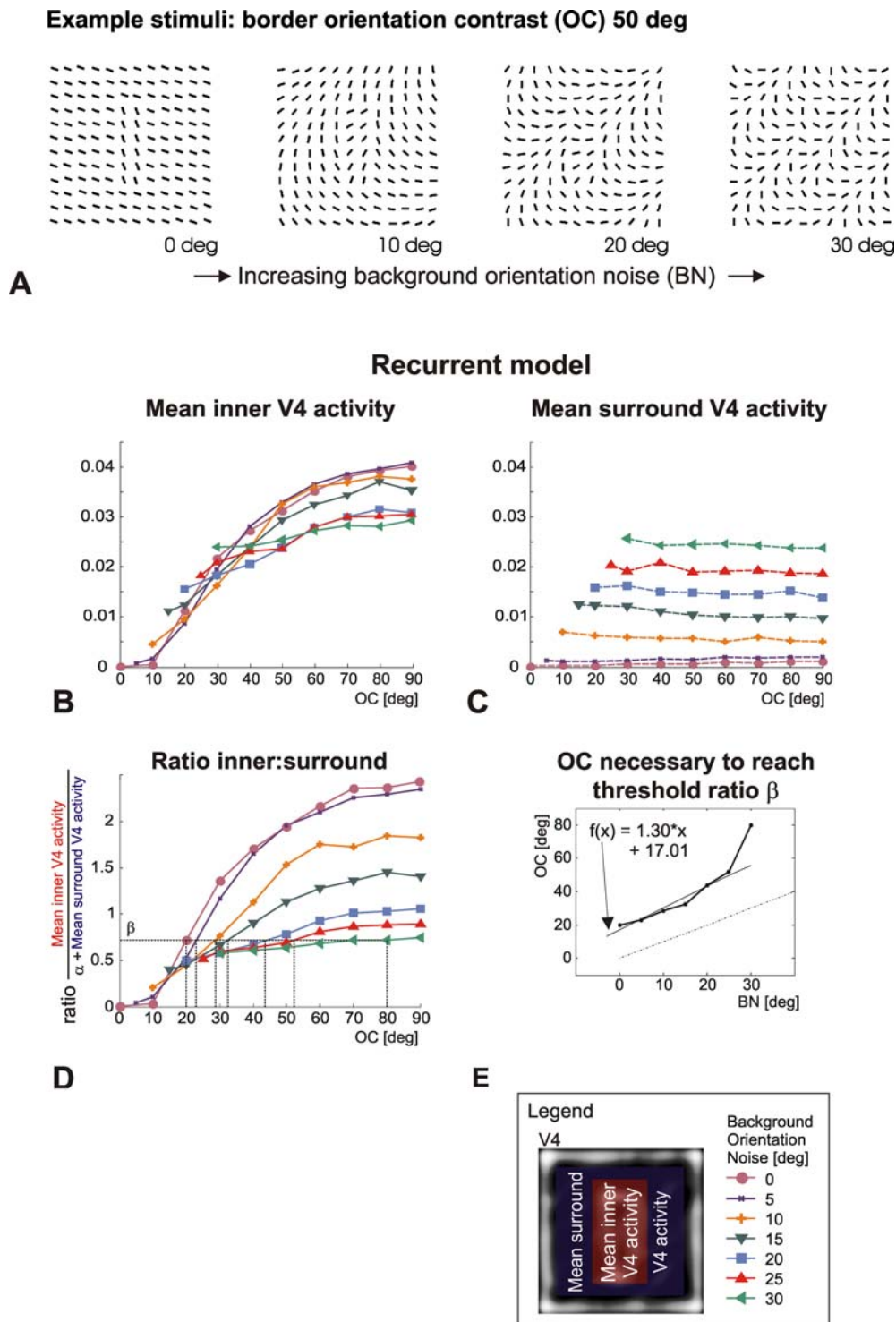


Fig. 6 Effects of a systematic variation of BN on model V4 activation patterns. (A) Example stimuli consisting of a pop-out bar defined by an OC of 50° embedded in various levels of orientation noise. The bar is visible for BN values up to 20°, but can hardly be detected for BN = 30°. (B) Mean model cell activity in the inner V4 region (marked as red rectangle), plotted in dependence of OC for various levels of BN. The mean inner V4 activity exhibits a sigmoidal dependency on OC for low levels of BN. However, the curves are markedly flatter for higher BN values. (C) Mean surround V4 activity (in the blue region), plotted in dependence on OC for various levels of BN. The level of surround

activity is almost independent from OC, but increases monotonically with increasing BN. (D) Ratio between mean inner and surround V4 activity. The curves are sigmoidal for low BN values, but get flatter with increasing BN. The OC necessary to reach the reference ratio β increases with increasing BN. (E) Threshold OC necessary for reference ratio β in dependence on BN. The curve is almost linear for BN values up to 25°. However, a very high OC is necessary for BN = 30°, indicating that the model’s ability to detect the bar gradually breaks down for orientation noise exceeding this level

independent of OC. This level exhibits a monotonic increase with increasing BN. In particular, for a BN of 30° , the mean inner V4 activity is only moderately enhanced above the surround activity level even for high values of OC (green lines with triangles in Fig. 6(B) and (C)). Consequently, strong background orientation noise results in an unspecific activation of the model V4 cells, which in turn prevents the formation of V4 activity to signal the border of the pop-out bar.

The ratios between the mean inner and surround V4 activities are plotted in Fig. 6(D). One can assume here that the texture border is detected by the model, when the ratio between inner and surround V4 activity exceeds a given threshold value β . Additionally, a constant (or tonic activation) α is added to the surround activity level. This prevents unspecific noise-related V4 activations in the inner region from causing high ratios in the case of low surround activity levels. For vanishing background orientation noise, the mean inner V4 activity significantly exceeds zero for OC values greater than 20° . In consequence, the ratio for $\text{BN} = 0^\circ$ and $\text{OC} = 20^\circ$ was taken as reference threshold value β (indicated as horizontal dashed line), and $\alpha = 0.0154$ was determined empirically so that the range of all curves was approximately centered on this value. For all values of BN a ratio larger than β , thus, indicates that the mean center activation exceeds the mean surround activation by a significantly amount. This is interpreted as an indicator that the model “detected” the pop-out bar.

The OC values for which the input-output curves reach the reference value β are monotonically shifted to the right with increasing BN. These threshold OC values are plotted against BN in Fig. 6(E). For BN up to 25° , the values increase approximately linearly. However, in case of $\text{BN} = 30^\circ$, very high OC value are necessary to reach the reference threshold β (see green line with triangles in Fig. 6(D)).

Taken together, in our model border orientation contrast in the input image results in V4 activity signaling the outline of the pop-out bar. The strength of this V4 activity exhibits a sigmoidal dependency on the OC in the input image. Background orientation noise, however, evokes unspecific V4 cell activations which in turn disturb the formation of V4 activity at the border of the pop-out bar. As a consequence, higher orientation contrasts are necessary to recover the models ability to detect the pop-out bar. For orientation noise exceeding approx. 30° , this ability gets gradually lost.

5.2.2 Feed-forward model

In a model without feedback connections, V4 cells lose their ability to modulate cell activity in the lower model areas V1 and V2. In the following, the impact of this limitation on the model performance in texture border processing was systematically tested by repeating the above depicted simu-

lations using a pure feed-forward model. The dependency of the mean inner V4 activity of the feed-forward model on the border orientation contrast is depicted in Fig. 7(A). As in the recurrent model, the input-output curves exhibit sigmoidal slopes for low values of BN. However, the initial rising flanks of these curves are less steep compared to the input-output curves of the recurrent model. Moreover, the flattening of the input-output curves with increasing BN is notably more pronounced than observed in the recurrent model. For BN values of 25° and 30° , the mean inner V4 activities do not increase at all with increasing OC (Fig. 7(A): lines with triangles pointing upward and to the left).

In Fig. 7(B) the mean V4 surround activity of the feed-forward model is depicted. Comparable to the fully recurrent model, the surround activity is more or less independent from OC for a specific value of BN and increases monotonically with increasing BN. However, for $\text{BN} = 25^\circ$ and 30° , the difference between the mean inner and surround activity is almost independent of the border OC (Fig. 7(A) and (B): lines with triangles pointing upward and to the left). Clearly, the feed-forward model was unable to detect the orientation contrast at the border of the pop-out bar in these cases even for high values of OC. Consequently, the curves depicting the ratios between inner and surround activity (Fig. 7(C)) remain approximately flat for BN values above 15° . Again, the ratio for $\text{BN} = 0^\circ$ and $\text{OC} = 20^\circ$ was taken as reference threshold value β ($\alpha = 0.0022$). When introducing background orientation noise, this reference value was reached by the feed-forward model only for BN values up to 15° (see Fig. 7(D)).

Taken together, the behavior of the feed-forward model was comparable to that of the recurrent model for low values of BN. However, the model’s ability to suppress unspecific activations evoked by orientation noise crucially depends on the modulation of cell activity in lower model areas via feedback from V4. Consequently, the feed-forward model’s performance in detecting the border of the pop-out bar breaks down for BN values above 15° . This in turn demonstrates that modulatory feedback enables the model visual system to achieve robustness of target detection over a wide range of stimulus conditions.

5.3 Discussion

The model replicates some of the key findings obtained in the psychophysical experiments of Nothdurft (1991, 1992). In his experiments subjects had to indicate the presence or absence of a pop-out bar while OC and BN were systematically varied. Correct response rates ranged between 50% (chance level) and 100%. His key observations were:

- For a BN of 0° , the curves indicating the number of correct responses (in%) exhibit a sigmoidal dependence on

Feedforward model

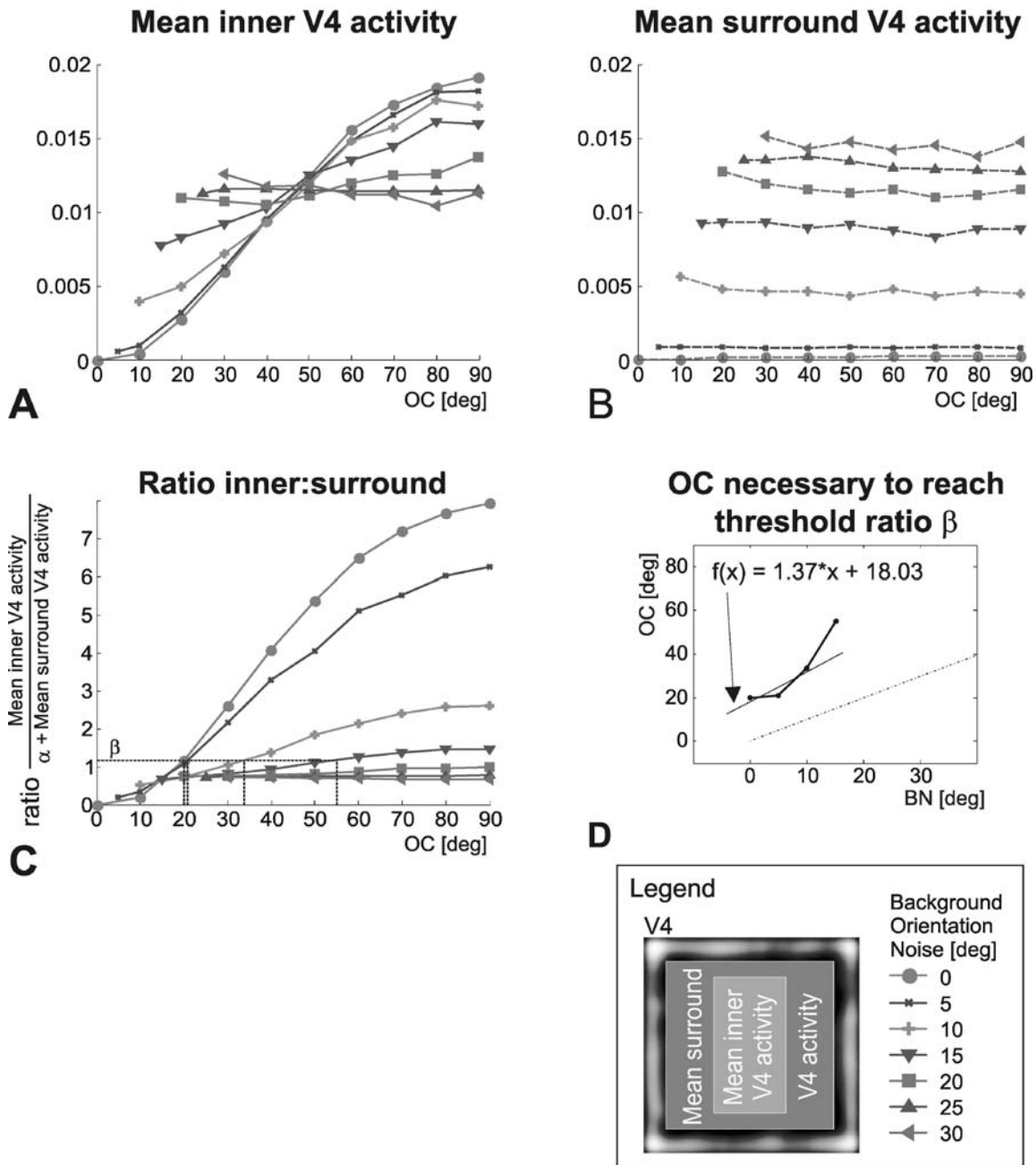


Fig. 7 Effects of background orientation noise on V4 activation patterns of a pure feed-forward model. (A) Dependency of mean inner V4 activity (light gray region) on OC for varying levels of BN. Compared to the fully recurrent model (Fig. 6), the curves get flat at lower values of BN. (B) Mean surround V4 activity (dark gray region) plotted in dependence on OC. It is almost independent from OC, but increases monotonically with increasing BN. For BN = 25° and BN = 30°, the surround activity is stronger than the inner V4 activity even for

high values of OC. (C) Ratio between mean inner and surround V4 activity. Compared to the fully recurrent model, the curves get flat at lower values of BN. Consequently, for BN values exceeding 20°, the reference ratio β is not reached even for the maximal OC of 90°. (D) Threshold OC necessary to reach reference ratio β in dependence on BN. The model’s ability to detect the texture bar breaks down already for BN values exceeding 15°

OC and approach 100% correct responses at $OC = 30^\circ$. Likewise, in our model, the ratio between inner and surround V4 activity shows a sigmoidal dependence on OC with its steepest increase for OC values up to 30° .

- In the psychophysical experiments, the correct response rates in dependence on OC are flattened with increasing BN. This is replicated by our model which has flatter ratios between the inner and surround V4 activity for higher values of BN.
- Nothdurft plotted the minimal OC necessary for a detection rate of 75% against BN. This curve increases almost linearly with increasing BN for values up to 20° – 25° . In our model, the minimal OC necessary to reach the reference ratio β exhibits an approximately linear dependence on BN for BN values up to 25° .
- Above a BN of 30° , subjects usually did not reach the 75% correct response level. Likewise, for $BN = 30^\circ$, the inner V4 activity level of our model is only marginally increased above the surround activity level even for high values of OC. This indicates that the model's ability to detect the orientation contrast at the border of the pop-out bar breaks down for BN values exceeding 30° .

Taken together, the V4 activation patterns of the full recurrent model exhibit exactly the same dependency on OC and BN as the detection performance of human observers. This allows tracing back the psychophysical results to a putative neural architecture. Furthermore, based on our simulation results, we propose that suppression of orientation noise crucially depends on the feedback connections between visual areas V1, V2 and V4.

6 General discussion

6.1 The proposed model and its key features

In this contribution we have proposed a recurrent model of texture boundary processing that builds upon recent knowledge of the structure and function of cortical architecture along the ventral pathway in primate vision. In particular, the model is mainly furnished to explain the contribution of different cortical areas that utilize simplified core mechanisms of spatial feedforward activity integration, lateral spatial excitatory and inhibitory interaction and modulatory feedback to re-enter activations from higher stages at early stages of spatially localized feature processing. Unlike previous modeling approaches that have favored more localized lateral long-range interactions based on intra-cortical processes, we have focused on mechanisms of inter-cortical processing that involves multiple cortical areas. Thus, the study provides computational evidence for the contributions of spatially distributed processing to implement robust surface boundary extraction under noisy conditions to reveal

camouflaged objects from backgrounds that have statistically similar properties. In particular, spatial density effects extensively studied parametrically in psychophysical experiments were explained quantitatively by the proposed model architecture. In that sense, the proposed model has previously been described to explain the detection properties of targets that were defined by oriented texture items. It was clarified how feature contrast effects contribute to the detection properties rather than the absolute detection of feature items. The computational processing properties have now been generalized showing how multiple feature dimensions, orientation and density, jointly contribute to object detection in cluttered scenes.

In our model, the positions of texture boundaries as well as pop-out target lines are detected via recurrent processing within a hierarchy of bi-directionally linked topographical areas. Model V4 cells use anisotropic receptive fields consisting of an excitatory center surrounded by two lateral inhibitory subfields to compare the orientation-selective input delivered by area V2. A line element evokes high V4 activity when it is *not* surrounded by approximately iso-oriented lines, independent whether it is a single target line embedded in a field of distractors or a line element belonging to the (continuous) border of a texture region. Consequently, model V4 cells gradually signal *local discontinuities* in element orientation and/or element density. The pooling of bottom-up input is followed by a process of intra-areal center-surround competition to sharpen the initial activation patterns and to normalize cell activities in a local neighborhood. The spatial distance at which neighboring line elements exert suppressive influence on each other is restricted by the V4 receptive field sizes and the range of center-surround competition, which both act in spatial regions of approximately equal size.

The resulting model V4 activity is finally fed back to modulate the initial model V2 activation pattern, enabling a recurrent flow of activity. Both the intra-areal center-surround competition and the process of feedback modulation help to enhance V4 activations due to salient pattern arrangements while at the same time suppressing ambiguous cell activities. This results in a context-dependent retuning of a cell's sensitivity to orientation discontinuities. Center-surround competition in model areas V1 and V2 acts in a spatial and orientational neighborhood. As a consequence, a stimulus that contains only minor background orientation noise will activate only specific orientation channels. In turn, this results in a weaker impact of center-surround competition on model V1 and V2 cell activation strength and in an increased sensitivity of model V4 cells to low orientation contrasts. In contrast, noisy stimuli activate a broad range of orientation channels, which compete with each other to attenuate the V2 bottom-up input to model area V4. In this case, higher orientation contrasts are necessary to evoke model V4 activity. Excitatory feedback from model area V4 results in an

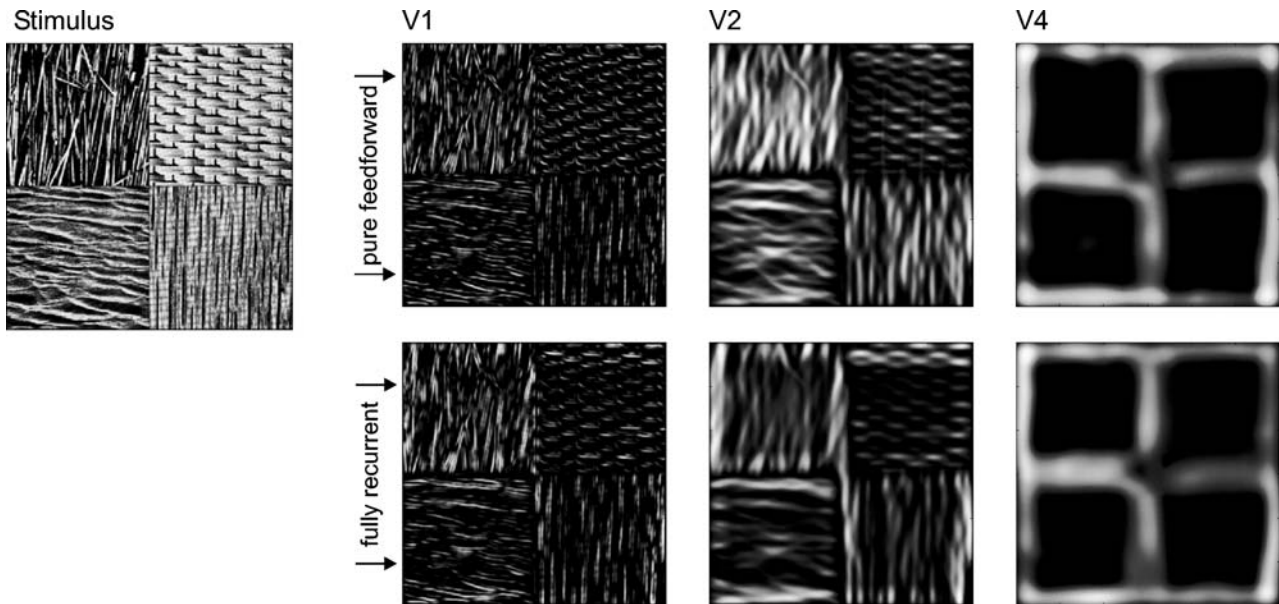


Fig. 8 Example of the model responses to four neighboring patches of oriented real-world structures (Brodatz, 1999: D15 Straw, D37 Water, D65 Handwoven Oriental rattan, D68 Wood grain; input image size 450×450). As the model V1 cells all work on the same fixed spatial scale, they do not succeed in capturing all details of the input patterns. However, the output of model area V1 is sufficient for area V4

to process the boundaries between the texture patches. As the orientation contrast between the patches is high, the V4 activation patterns of the feedforward and the recurrent model are quite similar. Maximal activation strength: $\max\{V1, V2, V4\} = \{0.043, 0.03, 0.057\}$ for the feedforward and $\{0.109, 0.098, 0.141\}$ for the recurrent model

orientationally and spatially specific enhancement of model V2 activity, in turn giving those V2 cells an advantage in the process of center-surround competition which signal putative “salient” element configurations. As demonstrated in chapter 5, this effect focuses the processing of the overall model on important image regions, increases the V4 cells’ sensitivity to salient orientation discontinuities and suppresses noise-related cell activities.

In Fig. 8 it is shown that the mechanisms outlined above also help to successfully segregate neighboring regions containing more complex real-world like oriented textures (Brodatz, 1999). In our model, the V1 simple and complex cells all work on the same fixed spatial scale. As a result, model V1 misses to represent some properties that appear at larger scales of the input texture patterns. However, as the model V4 cells integrate information over extended regions of the V1 output, area V4 is able to robustly signal the boundaries outlining the texture patches. Further examples for the successful interaction between V4 and the lower model areas in segregating more complex textures are demonstrated in Fig. 9. In Fig. 9(A) two inhomogeneous texture regions consisting of upright and tilted crosses are shown. The inhomogeneity is caused by randomly deleting 20% of the crosses and it is clearly signaled by the model V1 activation pattern. However, as model V4 cells integrate texture information over extended areas in the input image, they succeed in recovering the border between the two regions. In Fig. 9(B), the two texture regions both contain oblique lines of the same ori-

entation. As salient discontinuities are detected by the model V4 cells for each orientation channel separately, the overall model V4 stage is capable of clearly signaling the border between the two regions. In Fig. 9(C), both regions consist of left- and rightward tilted lines and differ only in the way the lines are arranged. The border between the two regions is recovered by a combination of model V2 and V4 mechanisms: At the border, several pairs of leftward tilted lines are grouped together by the model V2 long-range mechanisms. This results in a slightly enhanced V2 activation strength at these positions, which is detected by model area V4 and further enhanced by the recurrent flow of activity from V4 to V2. The model V4 activity signaling the border remains comparatively weak, which agrees well with the subjective impression of a somewhat fuzzy transition from one region to the other. In Fig. 9(D), the grouping of pairs of lines at the border between the two regions again results in an enhanced V2 activity which is further stabilized by recurrent V4–V2 interaction. In particular, Fig. 9(C) and (D) demonstrate how V2 grouping mechanisms and V4 mechanisms which signal salient activity transitions can interact in order to robustly process boundaries between texture regions.

6.2 Relation to psychophysics

The computational experiments presented in this paper allow to link several key results of the psychophysical studies of Nothdurft (1991, 1992, 2000c) on human texture

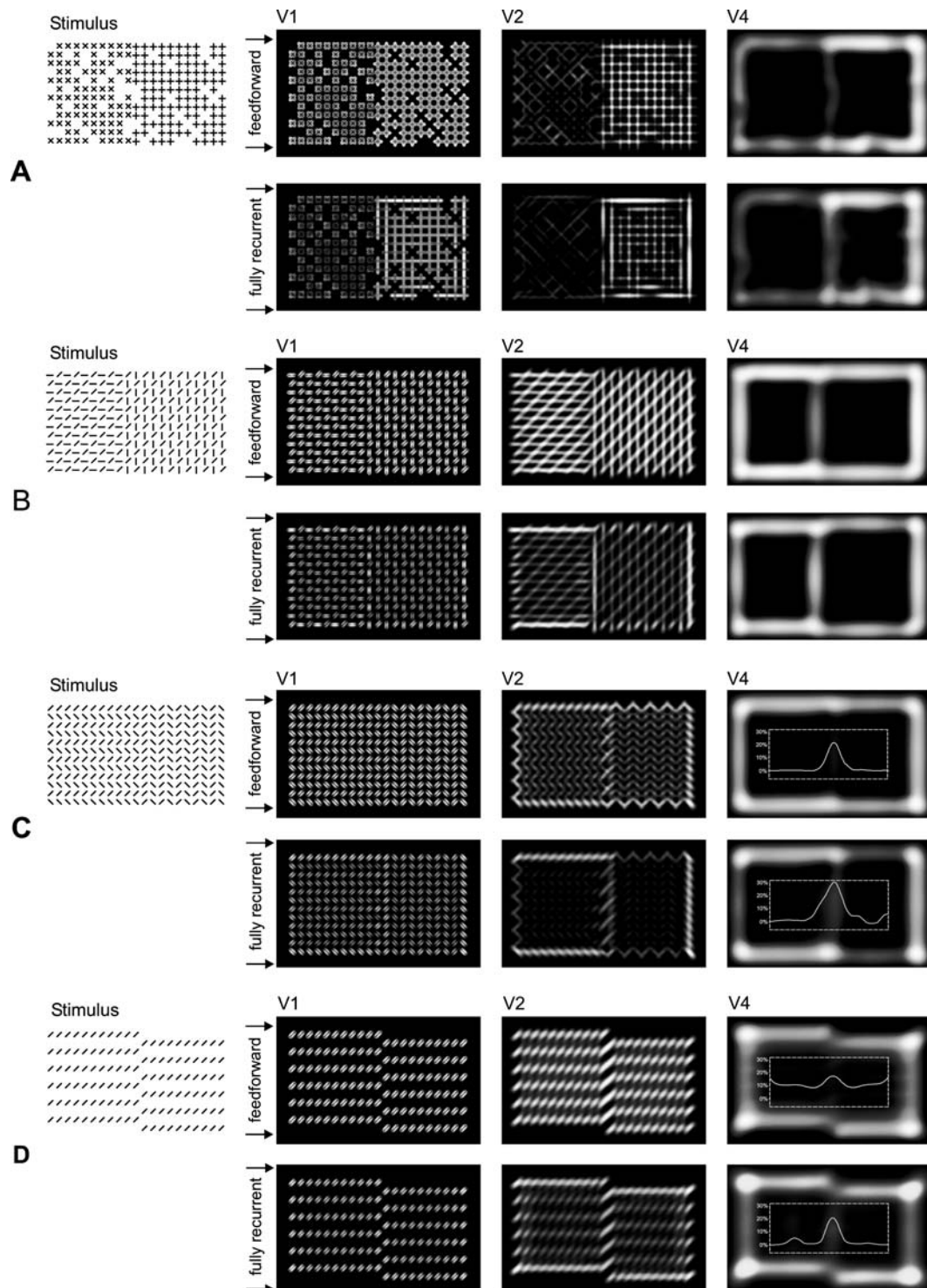


Fig. 9 Model responses to texture regions containing more complex artificial patterns (adopted from Li, 2000, Fig. 7). (A) Both texture patterns are inhomogeneous. Maximal activation strength of simulation: $\max\{V1, V2, V4\} = \{0.035, 0.028, 0.036\}$ (feedforward) and $\{0.098, 0.082, 0.091\}$ (recurrent). (B) Half of the line elements in both regions have the same orientation (rightward tilted). Maximal activation strength: $\max\{V1, V2, V4\} = \{0.041, 0.026, 0.044\}$ (feedforward) and $\{0.109, 0.084, 0.092\}$ (recurrent). (C) Both regions consist to identical proportions of left- and rightward tilted lines, which are arranged in two distinct patterns. Maximal activation strength: $\max\{V1, V2,$

$V4\} = \{0.046, 0.019, 0.043\}$ (feedforward) and $\{0.085, 0.051, 0.082\}$ (recurrent). (D) The complete texture consists of rightward tilted lines which are offset along the y-axis at the border between the two regions. Maximal activation strength: $\max\{V1, V2, V4\} = \{0.048, 0.024, 0.036\}$ (feedforward) and $\{0.097, 0.066, 0.070\}$ (recurrent). (C) and (D): In order to better visualize the rather weak activation patterns in the central part of area V4 (indicated by the dotted lines), the mean response strength was calculated in that region and overlaid to the original V4 activation pattern (scaling in % of the maximal V4 activation strength)

processing with activation patterns in area V4 of our model. We demonstrated that reading out model V4 activation patterns using simple center-surround mechanisms is sufficient to reproduce Nothdurft's findings on (i) effects of varying texture density, (ii) alignment effects and (iii) effects of background orientation noise. This enables us to trace back his results to the putative underlying neural mechanisms and to the relevant target areas. *First*, we demonstrated that the effects of texture density on the saliency of a pop-out target most likely depends on V4 receptive field sizes, the range of V4 intra-areal center-surround competition and a V4 receptive field organization of orientation selective cells having an excitatory center and lateral inhibitory surrounds (chapter 3). The increased saliency ratings of both pop-out and non-pop-out lines at low texture densities can be traced back to the limited receptive field sizes and range of center-surround competition. The peak saliency ratings of pop-out lines at medium texture densities depend on the receptive field organization. As previously discussed (Thielscher and Neumann, 2003), it is unlikely that, e.g. the anisotropic horizontal connections in V1 can account for these findings, as their integration range is limited and substantially smaller than the receptive fields in V4. This view is also supported by the findings of Nothdurft that the observed density effects substantially differ with feature modality (luminance vs. orientation vs. motion; Nothdurft, 2000c). In particular, the difference between the saliency curves of orientation- and motion-defined pop-out bars led him to suggest that “orientation and motion are processed in different subsequent areas (for example, in V4 and MT)”, i.e. not commonly in V1 but in different, higher visual areas.

In our simulation study, the V4 receptive field sizes are constant in order to reduce the complexity of the model. However, in the human visual maps the receptive fields monotonically increase in size from the representation of the fovea to the periphery. Consequently, the activation patterns of human V4 should exhibit a similar effect as observed above when a fixed density input is presented in the fovea and gradually moved to the periphery.

Second, we have demonstrated that the detectability of a pop-out bar (Nothdurft, 1992, p. 360) is influenced by the alignment of line items that compose the texture field. We suggest that this effect is most likely caused by the contributions of anisotropic mechanisms for contour integration and completion, as observed for V2 neurons (v.d. Heydt et al., 1984). In our model, the common grouping of both line elements that belong to the pop-out bar and those of the background to continuous “flow patterns” by model V2 bipole cells tends to reduce the orientation contrast “seen” by model V4 cells (chapter 4). In particular, model V4 activity is notably reduced when V2 bipole cells are able to group line elements across the long border of the pop-out bar, resulting in a less robust “detection” of the bar by our model. In addition, long-

range mechanisms for contour integration in V1 (mediated by anisotropic horizontal projections) might complement the contour effects observed in V2, to further increase and stabilize the alignment effect (Gilbert and Wiesel, 1989; Hirsch and Gilbert, 1991; Kapadia et al., 1995, 2000; Li, 2002; Hansen and Neumann, 2004). From a computational point of view, such V1 mechanisms are mainly a replication of the mechanisms implemented by our V2 bipole cells, merely acting on a reduced spatial scale. Consequently, they are not expected to change the overall model behavior, at least when considering the stimuli used in our simulations. We therefore decided not to integrate such V1 mechanisms in our model in order to reduce its overall complexity.

Third, we demonstrated that the increase of orientation contrast necessary to detect a pop-out bar embedded in a background of increasing orientation noise can be accounted by taking the ratio between inner model V4 activity (border of the bar) and surrounding V4 activity (orientation noise). The resulting quantitative measures were observed in psychophysical experiments (Nothdurft, 1991, 1992). This ratio has to reach a certain threshold in order to get a clear-cut model V4 activity signaling the border of a pop-out bar and in turn enabling its detection (chapter 5). In particular, the OC necessary to reach the threshold ratio increases almost linearly for BN values up to 25°. For higher values of BN, the model's ability to detect the border of the pop-out bar breaks down even for a maximal OC of 90°, thereby resembling the performance of human observers. As already outlined above, this behavior can be explained by (i) the increasing number of V1 and V2 orientation channels activated at increasing levels of orientation noise and their tendency to camouflage salient orientation discontinuities, and (ii) feedback from model area V4 in order to concentrate processing on “salient” image regions, thereby counteracting the effects of orientation noise for BN values up to 25° (see chapter 6.1 “Key model features”).

Taken together, our simulation study clearly supports Nothdurft's view that human texture processing is based on the detection of salient feature discontinuities by demonstrating that a biologically plausible implementation of this principle is capable to replicate several of his key findings. His results can be linked to the putative underlying neural mechanisms, target areas and key processing principles. This includes, e.g. V4 receptive field sizes and organization which determine the range of texture density effects. Furthermore, center-surround interactions and activity modulation via feedback connections is demonstrated to be crucial for the suppression of orientation noise.

Our simulation results on the effects of texture density also allow to interpret the findings of a recent psychophysical study (Meinecke and Donk, 2002) on visual search efficiency in texture fields of varying homogeneity. Stimulus homogeneity was controlled by replacing an increasing number of

the line elements of a homogeneous texture with blank gaps. In that study, it was demonstrated that the detection of a perpendicular target line embedded in an inhomogeneous field of iso-oriented distractors follows a U-shaped dependency on the number of line elements: Detection performance is high for stimuli containing only a few lines, decreases with increasing element number and starts to increase again for high element numbers. Based on our results, we suggest that pre-attentive segmentation processes are inefficient for low element numbers (i.e. low texture densities), resulting in (separate) localized V4 activity blobs for each line element. In this case, detection performance depends on the capability of subsequent attentive processes and scales inversely with the number of line elements. For high element numbers, however, pre-attentive texture border processing achieves segmentation of the stimulus and isolated V4 activity at the position of the target line enables efficient search and fast target detection (see also Thielscher et al., 2002). Our interpretation is supported by ERP results (Schubö et al., 2001; Schubö, 2002) which complement the psychophysical study by Meinecke and Donk (2002): Differences in posterior N2 amplitude (a brain potential specific to pre-attentive processing) between trials with and without a target line were only observed for high element numbers. This indicates that pre-attentive processing is only efficient when texture density is sufficiently high.

6.3 Relation to physiological and neuroimaging data

Physiological and neuroimaging studies indicate that visual area V4 is strongly involved in texture processing (de Weerd et al., 1996; Merigan, 1996, 2000; Kastner et al., 2000). For example, using fMRI, it was demonstrated that areas V4 and TEO are significantly stronger activated by texture borders than lower visual areas (Kastner et al., 2000). Likewise, V4 lesioned monkeys were unable to identify the orientation of a pop-out bar composed of several line items (i.e. they could not recover its outline defined by the texture border), although they still detected the presence of high OCs in the stimulus (Merigan, 2000). In our model, we propose a biologically plausible implementation of texture border processing by V4 cells, which traces back the sensitivity to salient orientation discontinuities to a putative underlying V4 receptive field organization having an excitatory center and two lateral inhibitory flanking regions. Key properties of our model V4 cells are supported by a study of Pollen et al. (2002) investigating the interaction of bar and grating stimuli in single cell V4 responses. The key observations in their study were that (i) individual neurons are predominantly selective to single input orientations and that (ii) RF obey a spatial center-surround organization. This indicates that V4 cells show stereotyped RF properties through selective integration of input activities from cells belonging to a

single specific orientation channel. These findings support the layout of our model V4 cells which (i) integrate input activations from specific orientation fields generated at previous model stages, and (ii) use a spatial center-surround organization of the RF to generate boundary activities at texture discontinuities. The view that at least a portion of V4 cells is selective to boundaries is also supported by an electrophysiological study of Pasupathy and Connor (2001). It was demonstrated that a significant amount of V4 cells selectively responds to boundaries of complex shapes defined by luminance transitions at specific positions in the stimuli.

A recent electrophysiological study demonstrates that saccades during a free viewing visual search task were directed toward the topographic locations of high V4 activity (Mazer and Gallant, 2003). For those cells being located around the target location of the next saccade an increased activity level was observed in approximately 50% of the recorded neurons. This increase was apparently determined solely by (bottom up) stimulus features. The activity of the other 50% of the neurons was additionally modulated by task demands and, thus, demonstrating the influence of feature-based attentional modulation (compare Reynolds et al., 2000). The authors concluded that “the spatial distribution of activity in V4 encodes the retinotopic locations of salient features throughout the visual field” (Mazer and Gallant, 2003, p. 1248). Restricted to the orientation domain, the V4 activation patterns of our model exactly replicate these findings: Model V4 activity is a gradual signal of the saliency of the underlying local orientation discontinuities in the input image, either evoked by a single target line or due to region borders. Although it was beyond the scope of the current modeling investigations, our general model architecture and dynamics allow to integrate task-dependent mechanisms, as observed in the above mentioned electrophysiological study. A feature-specific modulation of V4 cell activity could be achieved by biasing model V4 center-surround competition to, e.g. a specific orientation channel using modulatory feedback interactions from higher model areas which represent visual areas involved in guiding attention (Itti et al., 1998). Such an additional bias by top-down modulatory enhancement provides receiving cells an advantage in the subsequent competition of cell activity in space-orientation domain. Together, modulatory feedback enhancement and divisive inhibition resemble a biased-competition mechanism for feature enhancement and selection (Reynolds et al., 2000; see Bayerl and Neumann, 2005 for a related mechanism based on the proposed architecture employed for motion analysis). Taken together, our modeling approach is able to link the psychophysical results of Nothdurft (1991, 1992, 2000c) on target salience and detectability, as discussed above, with the electrophysiological results of Mazer and Gallant (2003) by proposing a putative neural implementation of salience

computations with the result being represented by model V4 activation patterns.

Several electrophysiological studies demonstrate that cells in visual area V1 respond to texture borders by exhibiting enhanced firing-rates (Gallant et al., 1995; Lamme et al., 1998, 1999; Nothdurft et al., 1999, 2000). Based on our modeling approach, we propose excitatory feedback from V4 cells vividly responding to the texture borders to be one of the neural origins of these modulatory effects. Feedback from V4 can act indirectly via area V2 to increase the V1 activity at corresponding topographical positions (Thielscher and Neumann, 2003).¹ Interestingly, this local increase results in a *decrease* of the mean V1 activity due to center-surround competition. When comparing mean activation levels for model variants with and without feedback connections (using, e.g. the texture density stimulus of line spacing 35, Fig. 4(A)), then the mean activity of model area V1 is reduced by 20% for the recurrent model compared to a pure feedforward model. Likewise, mean V2 activity is reduced by 44%. This model behavior is in accordance with findings by Murray et al. (2002) who used fMRI to demonstrate that V1 activity is reduced in response to coherent compared to random line drawings. This observation has been taken as evidence against prevalent theories that feedback exclusively enhances feedforward driven activities in resemblance to attentional selection effects (Koch and Poggio, 1999). Instead, a previous proposal that feedback may provide a signature to compare with the input (as in predictive coding schemes) to measure the discrepancy between higher-order cell expectancy and the stimulus responses at earlier stages (Rao and Ballard, 1999) have now gained experimental support. Murray et al. (2002) concluded that “higher areas may serve to disambiguate activity patterns in lower areas, reducing irrelevant activity and enhancing activity appropriate for the percept”. In our model, the net effect of feedback together with activity normalization results in a suppression of orientation noise and enhances those activities in areas V1 and V2 located at topographically corresponding positions as the predicting model V4 activity blobs. This mechanism of feedback integration allows to combine the coarse-grained detection of salient texture discontinuities with high-resolution localization properties at lower areas (see also Thielscher and Neumann, 2003, Fig. 9).

¹ Thielscher and Neumann (2003) also investigated direct feedforward/feedback interactions between localized model V1 and coarse-grained model V4 interaction in concert with and without the integrated action of model V2 to study the computational consequences of parametric variation of the relative strength of the V1–V4 coupling in comparison to V2–V4 coupling. Here, we did not conduct such an additional study for investigating the density effects since we wanted to keep the modeling as simple as possible already demonstrating the desired density effects in the strictly hierarchical V1–V2–V4 model.

The horizontal long-range projections within area V1 are proposed to provide a further, independent neural origin underlying the observed border enhancement effect (Gilbert and Wiesel, 1989; Hirsch and Gilbert, 1991; Li, 2002; Hansen and Neumann, 2004). Apart from the processing of luminance contrast V1 cell responses to texture patterns and texture boundaries have been investigated in several recent studies (Kapadia et al., 1995; Nothdurft et al., 2000). The latter investigation, in particular, aims at finding out whether a texture boundary defined by an orientation contrast in the arrangement of individual items is signaled by V1 cells. The key finding was that a cell located at a texture boundary that is defined by an abrupt change in bar orientation is enhanced in a modulatory fashion. In a similar vein, Lamme and coworkers (Zipser et al., 1996; Lamme et al., 1998, 1999) investigated the late response modulations of V1 cells to texture patterns that define a figural pattern that segregates against the background. Such response latencies of V1 cell receptive field modulations from outer-surround regions provide ample evidence for a processing mechanism that affects, or evaluates, responses generated during the feedforward sweep of initial stimulus response over several stages of hierarchical processing (Lamme and Roelfsema, 2000). Such late responses may in principle have different causes and the experiments mentioned do not address the specific underlying interactions from the surround pattern (Nothdurft et al., 2000). Lateral long-range interactions in V1 spanning several cortical hypercolumns might generate the observed response modulations. Localized patchy feedback connections that modulate the early cell responses may also generate a similar net effect though effectively spanning a broader spatial range due to the larger receptive field sizes of cells higher up in the processing hierarchy. Physiological studies demonstrate that the spatial field of activity integration in V1 is rather small in comparison with extrastriate cell receptive fields even when considering horizontal long-range connections. Cavanaugh et al. (2002a, b) calculated cortical magnification factors representing the receptive field sizes of V1 cells in terms of cortical distances in mm. The authors demonstrate that the majority of receptive field radii fell within a range of approximately 3 to 4 mm on the cortical map. Horizontal long-range connections in V1 span up to distances of 8 mm (Gilbert and Wiesel, 1989; Hirsch and Gilbert, 1991). In other words, the diameter of the field of activity integration via long-range connections is (roughly estimated) approximately 2 to 3 times the diameter of a cell’s receptive field. However, this is about 3 to 4 times smaller than V4 receptive fields (Smith et al., 2001).

The temporal evolution of the responses that reflect contextual effects might also help to gain further insights into the contributions of feedback and lateral response interactions. Girard et al. (2001) investigated the response latencies and

conductance velocities of connections between V1 and V2 to see how fast cells in one area can drive recipient cells. The authors found that the distributions of latencies were comparable for feedforward and feedback connections, with clear-cut peaks around 1.5 to 2 ms. Unlike this pattern, latencies of lateral horizontal responses distribute much broader with the majority of latencies ranging from 1.5 to 6 ms and the maximal latencies being up to 10 ms. These findings led the authors conclude that the main processing sweeps (feedforward and feedback) act fast and on mainly the same temporal scale. In particular, the impact of feedback modulation on cell activity in V1 can be expected to be as fast or even faster than the activity modulation by neighboring V1 cells mediated by the horizontal V1 long-range connections. It is not unrealistic, therefore, to interpret these findings as indication for a tight coupling of feedforward and feedback processing to fuse localized feature measurements with context information that utilizes feature integration in different conjunctions and contrast measurements on a much broader scale. The late responses observed by, e.g. Lamme's investigations are then attributed to the polysynaptic stages of feedforward processing and integration and the subsequently initiated feedback projection from higher to lower stages of processing.

Our model proposal draws upon these findings. Although we do not rule out contributions from additional lateral interactions based on intra-cortical long-range interactions, our focus is on feedforward and feedback processing that implement different modes of vision in feature-driven forward and modulatory recurrent processing (Lamme and Roelfsema, 2000). We primarily investigated and provided computational evidence that individual oriented feature items and their local orientation contrast is signaled during the cortical feedforward sweep. Feedback that is generated at various stages higher up in the processing hierarchy provides a reentry signal to spatially modulate the initial responses and thus combine localized signals with coarse-grained context patterns which encode salient contrast patterns.

6.4 Other models

Our modeling approach is related to and shares properties of (i) models of recurrent V1–V2 interaction and (ii) models of human texture processing.

Models of recurrent V1–V2 interaction simulate processes of robust contour formation observed in early vision (Grossberg and Mingolla, 1985; Neumann and Sepp, 1999; Mansson, 2000; Finkel and Edelman, 1989; Parent and Zucker, 1989; Finkel and Sajda, 1992; Heitger et al., 1998). They can reproduce the induction of illusory contours, as observed in V2 contour neurons (v.d. Heydt et al., 1984), as well as the spatial grouping of texture items and the high saliency

of aligned bars. The computational capabilities of these models are preserved and incorporated by our approach, resulting in a model being in accordance with a more general architecture of form processing in the ventral visual stream. In particular, our model demonstrates how grouping of aligned contour items by V1–V2 interaction influences the capabilities of the overall model in texture boundary detection.

Many models trying to elucidate the neural mechanisms which underlie human texture processing propose a two-stage strategy (e.g., Malik and Perona, 1990; Bergen, 1991; Landy and Bergen, 1991; Graham et al., 1993; Kehrner and Meinecke, 2003): First, the input image is analysed using a set of filters having different orientations and (in some cases) different spatial frequencies. Subsequently, a second processing stage detects activity changes in the output of the first filter bank in order to establish borders between different texture regions. These approaches resemble the scheme of texture segregation of Beck (1982) who proposed a hierarchy of several processing levels, namely feature detection, linking, difference encoding and (threshold-based) decision for segmentation. These strategies are comparable to the hierarchy of computation implemented by our feed-forward stream of model V1 cells (filtering the input image) followed by model V4 cells (detecting changes in the V1 activity patterns relayed via V2). However, unlike our model, these approaches cannot explain the psychophysical findings of Nothdurft, as demonstrated in this study. They do not elucidate either the neural origins of alignment effects or the putative role of feedback from higher model areas. As demonstrated in this study, a pure feed-forward processing hierarchy consisting of successively repeating stages of pooling of bottom-up input followed by center-surround competition results in a less robust identification of texture borders in the presence of orientation noise compared to a model with additional modulatory feedback connections (chapter 5). Consequently, only the model with intact feedback can account for the performance of human observers in the detection of texture bars. A detailed discussion of the relation of our approach to other models of texture processing can be found in Thielscher and Neumann (2003).

A model of human texture boundary detection segregation has been proposed by Li (2000) suggesting that spatially anisotropic cooperative/competitive interactions in area V1 alone can explain a variety of texture data and stimuli used in visual search experiments (Li, 2002). The core element of the model consists of utilizing spatially anisotropic kernels which integrate input from similar oriented cells in spatially separate sub-fields in the surround of a target cell. Here excitatory input is sampled from regions in co-axial direction of the target cell orientation, while inhibitory input is generated from regions orthogonal to the target cell orientation, utilizing weighting functions both displaying a figure-eight, or bow-tie, shape. The excitatory kernel allows for long-range

integration of those activities which are aligned along the cell's axis of orientation preference, similar to the model of recurrent V1–V2 interaction proposed by Grossberg and Mingolla (1985) or in the V1 processing stage of Grossberg and Raizada (2000). A key difference appears in the definition of the inhibitory mechanism between those models and the one proposed in Li (2000). In the latter, excitatory and inhibitory inputs are defined such that in case of stimulation with a field of equally oriented bars the central target cell receives stronger inhibition than excitation. As a consequence the cell's response to a bar that appears as part of a texture arrangement is inhibited and, thus, its activation is lowered. At the location of a region border, where the homogeneity in the arrangement of the items is disrupted, a target cell now receives only minor inhibitory input, which, in turn, causes less reduction in cell activity in response to an item located at the boundary. The relative difference between responses at texture borders and within homogeneous textures leads to an increase of saliency measure at boundaries. In sum, the functionality of texture boundary detection is mainly influenced by the contribution of the inhibitory mechanism in this model. In other models, like those of Grossberg and coworkers, inhibition is crucial for enabling the competition between neighboring cells to enhance those activations representing salient items in the image. In our model, a similar behavior is accomplished by utilizing shunting inhibition to normalize the target activities (supported by collinear activities from the excitatory sub-field) against their surround. This renders inhibition also of central importance to enhance activities at border locations that were receiving excitatory input from coherently aligned and equally oriented items.

One simulation study reported by Li (2000) investigates effects of orientation noise, texture bar spacing and texture element alignment generated by recurrent V1 interactions. When probing her model with stimuli of different texture densities, V1 activity correlates with human detection performance (Nothdurft, 1985) up to bar spacing/bar length ratio of 3, but quickly deteriorates for higher ratios (Li, 2000, Fig. 6(K)). In contrast, human detection performance and the performance of our model which involves area V4 show a slower and smoother decline with decreasing texture density. This indicates that in case of texture patterns of low density the spatial extent of long-range interaction within Li's model of V1 is limited compared to the range of spatial integration of texture information by the human visual system. We suggest that the observed behavior may arise from a combination of cell responses in V1 with the reentry signal generated by cells in higher visual areas that have larger receptive fields and can thus provide a much broader contextual view of a scene. It is the context information that we suggest is delivered robustly by higher-order centers along the cortical processing pathways. We propose that V1 horizontal long-range projections

implement a powerful multi-purpose preprocessing mechanism that enhances the signal-to-noise ratio (particularly for low-contrast stimuli) and may enhance multi-dimensional feature patterns such as corners and localized junctions (see discussion in Hansen and Neumann, 2004).

The robust identification of texture borders crucially depends on the reliable identification of meaningful orientation discontinuities and may, therefore, be improved by more flexible mechanisms such as provided by the feedforward and reentry mechanisms proposed in our model. With our contribution, we demonstrate that alignment effects mediated by anisotropic connections in model area V2 (which have comparable layout to those used in Li's model of V1 integration cells) affect model V4 activity to an extent which is similar to the one observed in human behavior. We suggest that the model proposed by Li (2000) cannot fully develop its strength for certain item configurations. For example, consider the stimulus configuration used in Fig. 5 of Li (2000). Here two homogeneous regions of equally oriented bars are spatially juxtaposed horizontally. Above threshold output at the border where the two regions meet is generated for those bars oriented parallel to the region boundary where cells receive full excitatory input but only half inhibitory input. For the region that consists of bars oriented orthogonal to the region border boundary cells get an approximately 50% reduction from both excitatory as well as inhibitory input. Thus, no saliency is generated on this side of the contour. We suggest that with the contribution of a contrast mechanism as the one proposed in our model, the enhancement at such boundaries is more balanced and, thus, stabilize the texture boundary detection process.

Furthermore, when considering the stimulus variations discussed above, the reliable formation of texture boundaries requires integration utilizing a larger neighborhood in the space and orientation domain. Based on several empirical findings, we have argued that the robust formation of texture boundaries is likely a major function of the processing in higher cortical areas. For example, the fMRI study of Kastner et al. (2000) demonstrates that texture segregation tasks cause significant activation in areas V1, V2 and V4 with an almost linear increase of hemodynamic activity along the feedforward hierarchy indicative for a significant role of the anatomically higher areas. In addition, response modulations that bridge large spatial ranges that reach beyond V1 lateral integration ranges provide further evidence for the role of feedback. In functional terms, feedback may provide a reentry signal (Finkel and Edelman, 1989) that modulates initially generated rough-and-ready estimates along the feedforward processing cascade to evaluate and partially enhance those signal contributions that match the more coarse-grained response generated at later stages fusing multiple feature channels and their contrast measures.

6.5 Conclusion

The goal of this study was to trace back several key results on human performance in texture processing, as obtained in psychophysical studies (Nothdurft, 1991, 1992, 2000c; Meinecke and Donk, 2002) to their neural origins. We systematically investigated the behavior of a biologically plausible model of texture border processing, which is based on a large database of physiological and anatomical data, to tightly link its activation patterns with the results of the psychophysical studies. This allowed us to identify key processing principles as well as key areas which putatively underlie human performance. In particular, it was demonstrated that alignment effects in human texture processing are likely to result from mechanisms of contour groupings in area V2. Density effects in texture processing probably reflect a specific V4 receptive field organization and the range of V4 intra-areal center-surround competition. Furthermore, we demonstrated that

Appendix A Model V2 bipole cells: spatial subfields

The weighting functions $K^{left/right}$ determining the spatial layout of the subfields are modeled as anisotropic Gaussians, which are cut off in the central part of the cell by means of a sigmoid function. The partial overlap of the subfields in the center of the cell defines the classical receptive field:

$$K_{i\theta}^{left} = \Lambda_{\sigma_{k-x}, \sigma_{k-y}, \tau_{k-x}, 0, \theta}(\vec{x}_i) \cdot \frac{1}{1 + \exp\left(-A_k \vec{x}_i \cdot \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} - B_k\right)}$$

$$K_{i\theta}^{right} = \Lambda_{\sigma_{k-x}, \sigma_{k-y}, \tau_{k-x}, 0, \theta}(\vec{x}_i) \cdot \frac{1}{1 + \exp\left(+A_k \vec{x}_i \cdot \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} - B_k\right)}$$
(9)

\vec{x}_i : Cartesian coordinates of point i

$\sigma_{k-x} = 22.0$; $\sigma_{k-y} = 1.0$; $\tau_{k-x} = 2.0$; $A_k = 0.8$; $B_k = 9.0$

Table A.1 Parameters of the three-stage cell dynamics and standard deviations σ of the Gaussians pooling the center and surround activity in the spatial

	Parameters for $l^{(1)}$				Parameters for $l^{(2)}$				σ of			
	α_1	β_1	γ_1	c_1	α_2	β_2	γ_2	δ_2	α_2	ψ^+	Δ^+	ψ^-
V1	12.0	0.73	3.7	25	1.0	2.8	3.5	5.0	0.2	1.0	2.0	3.0
V2	12.0	0.85	4.2	20	1.0	2.9	3.1	50.0	0.2	2.0	2.0	6.0
V4	–	–	–	–	1.0	10.6	9.9	1000	0.2	8.0	2.0	24.0

and orientational domains (see Eqs. (10) and (11)). The values were kept constant through all experiments

noise suppression is probably the main functional role of the dense feedback connections from V4 to V2 in the context of texture processing. In particular, feedback from V4 was identified to be crucial in order to raise the model performance to the level of human observers. In the model, orientation noise activates adjacent orientation channels and thereby camouflages important orientation discontinuities signaling, e.g. a region border. This effect is counteracted by mechanisms of center-surround competition in model areas V2 and V4 as well as feedback from V4 in order to suppress ambiguous cell activities and to focus the processing of the overall model on salient discontinuities. In behavioral terms, such influences from recurrent cycles of processing help to detect objects even when they are camouflaged. Therefore, these objects were rendered visible when there is enough time for detection (Lamme and Roelfsema, 2000). The gradual cell activation dynamics implemented in our model enables area V4 to act as salience map in the orientation domain, thereby signaling the position of putative important image regions in a graded manner.

Acknowledgments We thank the reviewers for their constructive criticism and helpful suggestions to significantly clarify and improve the manuscript.

References

Bayerl P, Neumann H (2005) Feature-based attention increases the selectivity of population responses in a model of the primate visual cortex. In: Proceedings of the 6th Int'l Workshop Neural Coding 2005, Marburg (Germany), pp. 55–56.

Beck J (1982) Textural segmentation. In: J Beck, ed. Organization and representation in perception. Lawrence Erlbaum Associates, Hillsdale, pp. 285–317.

Bergen JR (1991) Theories of visual texture perception. In: D Regan, ed. Spatial vision, vol 10. Macmillan Press, pp. 115–134.

Brodatz P (1999) Textures: A Photographic Album for Artists and Designers. Dover Publications.

Cavanaugh JR, Bair W, Movshon JA (2002a) Nature and interaction of signals from the receptive field center and surround in Macaque V1 neurons. J. Neurophysiol. 88: 2530–2546.

Cavanaugh JR, Bair W, Movshon, JA (2002b) Selectivity and spatial distribution of signals from the receptive field surround in Macaque V1 neurons. J. Neurophysiol. 88: 2547–2556.

- Crick F, Koch C (1998) Constraints on cortical and thalamic projections: the no-strong-loops hypothesis. *Nature* 391: 245–250.
- de Weerd P, Desimone R, Ungerleider LG (1996) Cue-dependent deficits in grating orientation discrimination after V4 lesions in macaques. *Vis. Neurosci.* 13: 529–538.
- Eckhorn R, Reitboeck HJ, Arndt M, Dicke P (1990) Feature linking via synchronization among distributed assemblies: Simulation of results from cat visual cortex. *Neural. Comput.* 2: 293–307.
- Felleman DJ, van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex* 1: 1–47.
- Finkel LH, Edelman GM (1989) Integration of distributed cortical systems by reentry: a computer simulation of interactive functionally segregated visual areas. *J. Neurosci.* 9: 3188–3208.
- Finkel LH, Sajda P (1992) Object discrimination based on depth-from-occlusion. *Neural. Comput.* 4: 901–921.
- Gallant JL, van Essen DC, Nothdurft HC (1995) Two-dimensional and three-dimensional texture processing in visual cortex of the Macaque monkey. In: TV Pappathomas ed. *Early Vision and Beyond*. The MIT Press, Cambridge.
- Gilbert CD, Wiesel TN (1989) Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J. Neurosci.* 9: 2432–2442.
- Girard P, Hupe JM, Bullier J (2001) Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities. *J. Neurophysiol.* 85: 1328–1331.
- Graham N, Sutter A, Venkatesan C (1993) Spatial-frequency- and orientation-selectivity of simple and complex channels in region segregation. *Vis. Res.* 33: 1893–1911.
- Grossberg S (1980) How does a brain build a cognitive code? *Psychol. Rev.* 87: 1–51.
- Grossberg S, Mingolla E (1985) Neural dynamics of perceptual grouping: textures, boundaries, and emergent segmentations. *Percept. Psychophys.* 38: 141–171.
- Grossberg S, Raizada RDS (2000) Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vis. Res.* 40: 1413–1432.
- Hansen T, Neumann H (2004) Neural mechanisms for the robust representation of junctions. *Neural Comput.* 16: 1013–1037.
- Heitger F, v.d. Heydt R, Peterhans E, Rosenthaler L, Kübler O (1998) Simulation of neural contour mechanisms: Representing anomalous contours. *Image Vis. Comput.* 6: 407–421.
- Hirsch JA, Gilbert CD (1991) Synaptic physiology of horizontal connections in the cat's visual cortex. *J. Neurosci.* 11: 1800–1809.
- Hubel DH, Wiesel TN (1959) Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* 148: 574–591.
- Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* 160: 106–154.
- Hupe JM, James AC, Payne BR, Lomber SG, Girard P, Bullier J (1998) Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature* 394: 784–787.
- Itti L, Koch C, Niebur E (1998) A model of saliency-based visual attention for rapid analysis. *IEEE PAMI* 20: 1254–1259.
- Kapadia MK, Ito M, Gilbert CD, 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: 843–856.
- Kapadia MK, Westheimer G, Gilbert CD (2000) Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *J. Neurophysiol.* 84: 2048–2062.
- Kastner S, de Weerd P, Pinsk MA, Elizondo MI, Desimone R, Ungerleider LG (2001) Modulation of sensory suppression: Implications for receptive field sizes in the human visual cortex. *J. Neurophysiol.* 86: 1398–1411.
- Kastner S, de Weerd P, Ungerleider LG (2000) Texture segregation in the human visual cortex: A functional MRI study. *J. Neurophysiol.* 83: 2453–2457.
- Kehrer L, Meinecke C (2003) A space-variant filter model of texture segregation: parameter adjustment guided by psychophysical data. *Biol. Cybern.* 88: 183–200.
- Koch C, Poggio T (1999) Predicting the visual world: Silence is golden. *Nat. Neurosc.* 2: 9–10.
- Krauskopf J (1963) Effect of retinal image stabilization on the appearance of heterochromatic targets. *J. Opt. Soc. Am.* 53: 741–744.
- Lamme V, Rodriguez-Rodriguez V, Spekreijse H (1999) Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the Macaque monkey. *Cerebral Cortex* 9: 406–413.
- Lamme V, Super H, Spekreijse H (1998) Feedforward, horizontal, and feedback processing in the visual cortex. *Curr. Opin. Neurobiol.* 8: 529–535.
- Lamme VAF, Roelfsema PR (2000) The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosci.* 2000: 571–579.
- Landy MS, Bergen JR (1991) Texture segregation and orientation gradient. *Vis. Res.* 31: 679–691.
- Li Z (2000) Pre-attentive segmentation in the primary visual cortex. *Spat. Vis.* 13: 25–50.
- Li Z (2002) A saliency map in primary visual cortex. *Trends in Cogn. Sci.* 6: 9–16.
- Malik J, Perona P (1990) Preattentive texture discrimination with early vision mechanisms. *J. Opt. Soc. Am. A* 7: 923–932.
- Mansson J (2000) Occluding contours: A computational model of suppressive mechanisms in human contour perception, vol. 81. Lund University Cognitive Studies, Lund.
- Martinez-Conde S, Macknik SL, Hubel DH (2004) The role of fixational eye movements in visual perception. *Nat. Rev. Neurosci.* 5: 229–240.
- Mazer JA, Gallant JL (2003) Goal-related activity in V4 during free viewing visual search: Evidence for a ventral stream visual salience map. *Neuron* 40: 1241–1250.
- Meinecke C, Donk M (2002) Detection performance in pop-out tasks: Nonmonotonic changes with display size and eccentricity. *Perception* 31: 591–602.
- Merigan WH (1996) Basic visual capacities in shape discrimination after lesions of extrastriate area V4 in macaques. *Vis. Neurosci.* 13: 51–60.
- Merigan WH (2000) Cortical area V4 is critical for certain texture discriminations, but this effect is not dependent on attention. *Vis. Neurosci.* 17: 949–958.
- Mignard M, Malpeli JG (1991) Paths of information flow through visual cortex. *Science* 251: 1249–1251.
- Mumford DB (1994) Neuronal Architectures for Pattern-theoretic Problems. In: C Koch, J Davis, eds. *Large-Scale Neuronal Theories of the Brain*. MIT Press, pp. 125–152.
- Murray SO, Kersten D, Olshausen BA, Schrater P, Woods DL (2002) Shape perception reduces activity in human primary visual cortex. *PNAS* 99: 15164–15169.
- Neumann H, Mingolla E (2001) Computational neural models of spatial integration in perceptual grouping. In: TF Shipley, PJ Kellman, eds. *From Fragments to Objects—Segmentation and Grouping in Vision*. Elsevier, Amsterdam, pp. 353–400.
- Neumann H, Pessoa L, Hansen T (1999) Interaction of ON and OFF pathways for visual contrast measurement. *Biol. Cybern.* 81: 515–532.
- Neumann H, Sepp W (1999) Recurrent V1–V2 interaction in early visual boundary processing. *Biol. Cybern.* 81: 425–444.
- Nothdurft HC (1985) Sensitivity for structure gradient in texture discrimination tasks. *Vis. Res.* 25: 1957–1968.
- Nothdurft HC (1991) Texture segmentation and pop-out from orientation contrast. *Vis. Res.* 31: 1073–1078.

- Nothdurft HC (1992) Feature analysis and the role of similarity in preattentive vision. *Percept. Psychophys.* 52: 355–375.
- Nothdurft HC (2000c) Saliency from feature contrast: Variations with texture density. *Vis. Res.* 40: 3181–3200.
- Nothdurft HC, Gallant JL, van Essen DC (1999) Response modulation by texture surround in primate area V1: correlates of “popout” under anesthesia. *Vis. Neurosci.* 16: 15–34.
- Nothdurft HC, Gallant JL, van Essen DC (2000) Response profiles to texture border patterns in V1. *Vis. Neurosci.* 17: 421–436.
- Parent P, Zucker S (1989) Trace inference, curvature consistency, curve detection. *IEEE PAMI* 11: 823–839.
- Pasupathy A, Connor CE (2001) Shape representation in area V4: Position-specific tuning for boundary conformation. *J. Neurosci.* 21: 2505–2519.
- Peterhans E (1997) Functional organization of area V2 in the awake monkey. In: KS Rockland, JH Kaas, A Peters, eds. *Extrastriate Cortex in Primates*, vol 12. Plenum Press, New York.
- Pollen DA, Przybyszewski AW, Rubin MA, Foote W (2002) Spatial receptive field organization of macaque V4 neurons. *Cerebral Cortex* 12: 601–616.
- Przybyszewski AW, Gaska JP, Foote W, Pollen DA (2000) Striate Cortex increases contrast gain of macaque LGN neurons. *Vis. Neurosci.* 17: 485–494.
- Rao RPN, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects.
- Reynolds JH, Pasternak T, Desimone R (2000) Attention increases sensitivity of V4 neurons. *Neuron* 26: 703–714.
- Ross WD, Grossberg S, Mingolla E (2000) Visual cortical mechanisms of perceptual grouping: Interacting layers, networks, columns, and maps. *Neural Netw.* 13: 571–588.
- Safran AB, Landis T (1998) The vanishing of the Sun. A manifestation of plasticity in the visual cortex. *Surv. Ophthalmol.* 42: 449–452.
- Salin PA, Bullier J (1995) Corticocortical connections in the visual system: structure and function. *Physiol. Rev.* 75: 107–154.
- Sandell JH, Schiller PH (1982) Effect of cooling area 18 on striate cortex cells in the squirrel monkey. *J. Neurophysiol.* 48: 38–48.
- Schubö A (2002) Is preattentive processing in visual search similar to preattentive processing in texture segmentation? In: HH Bülthoff, KR Gegenfurtner, HA Mallot, R Ulrich eds. *TWK. Knirsch Verlag, Tübingen*, p. 191.
- Schubö A, Meinecke C, Schröger E (2001) Automaticity and attention: investigating automatic processing in texture segmentation with event-related brain potentials. *Cogn. Brain Res.* 11: 341–361.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RBH (1995) Borders of multiple visual areas in humans revealed by functional MRI. *Science* 268: 889–893.
- Smith AT, Singh KD, Williams AL, Greenlee MW (2001) Estimating receptive field size from fMRI data in human striate and extrastriate visual cortex. *Cerebral Cortex* 11: 1182–1190.
- Thielscher A, Neumann H (2003) Neural mechanisms of cortico-cortical interaction in texture boundary detection: A modeling approach. *Neuroscience* 122: 921–939.
- Thielscher A, Schubö A, Neumann H (2002) A neural model of human texture processing: Texture segmentation vs. Visual search. In: HH Bülthoff, C Walraven eds. *Biologically Motivated Computer Vision*. Springer, Heidelberg.
- Ungerleider LG, Haxby JV (1994) ‘What’ and ‘where’ in the human brain. *Curr. Opin. Neurobiol.* 4: 157–165.
- v.d. Heydt R, Heitger F, Peterhans E (1993) Perception of occluding contours: Neural mechanisms and a computational model. *Biomed. Res.* 14: 1–6.
- v.d. Heydt R, Peterhans E, Baumgartner G (1984) Illusory contours and cortical neuron responses. *Science* 224: 1260–1262.
- Wolfson SS, Landy MS (1995) Discrimination of orientation-defined texture edges. *Vis. Res.* 35: 2863–2877.
- Yarbus AL (1967) *Eye movements and vision*. Plenum Press, New York.
- Zilles K, Clarke S (1997) Architecture, connectivity, and transmitter receptors of human extrastriate visual cortex—Comparison with nonhuman primates. In: KS Rockland, JH Kaas, A Peters eds. *Extrastriate Cortex in Primates*, vol 12. Plenum Press, New York.
- Zipser K, Lamme VAF, Schiller PH (1996) Contextual modulation in primary visual cortex. *J. Neurosci.* 16: 7376–7389.