Biol. Cybern. 88, 236–246 (2003)
DOI 10.1007/s00422-002-0378-2

Biological Cybernetics

Suppression of contour perception by band-limited noise and its relation to nonclassical receptive field inhibition

Nicolai Petkov, Michel A. Westenberg

Institute of Mathematics and Computing Science, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands

Received: 11 July 2002 / Accepted in revised form: 19 November 2002 / Published online: 28 February 2003

Abstract. Band-spectrum noise has been shown to suppress the visual perception of printed letters. The suppression exhibits a specific dependence on the spatial frequency of the noise, and the frequency domain of most effective inhibition has been related to the size of the letters. In this paper, we address two important questions that were left open by previous studies: (1) Is the observed effect specific to text, and which parameters determine the domain of most effective suppression? (2) What is the origin of the effect in terms of underlying neural processes? We conduct a series of psychophysical experiments that demonstrate that the frequency domain of most effective inhibition depends on the stroke width of the letter rather than on the letter size. These experiments also demonstrate that the effect is not specific to the recognition of letters but also applies to other objects and even to single bars. We attribute the observed effect to nonclassical receptive field (non-CRF) inhibition in visual area V1. This mechanism has previously been suggested as the possible origin of various other perceptual effects. We introduce computational models of two types of cell that incorporate non-CRF inhibition, which are based on Gabor energy filters extended by surround suppression of two kinds: isotropic and anisotropic. The computational models confirm previous qualitative explanations of perceptual effects, such as orientation contrast pop-out, reduced saliency of lines embedded in gratings, and reduced saliency of contours surrounded by textures. We apply the computational models to the images used in the psychophysical experiments. The computational results show a dependence of the inhibition effect on the spatial frequency of the noise that is similar to the suppression effect measured in the psychophysical experiments. The experimental results and their explanation give further support to the idea of a possible functional role of non-CRF inhibition in the separation of contour from texture information and the mediation of object contours to higher cortical areas.

Correspondence to: N. Petkov
(e-mail: N.Petkov@cs.rug.nl)

1 Introduction

Band-spectrum noise has been demonstrated to suppress the detection and recognition of printed letters: the luminance contrast, above which letters are correctly recognized, has been shown to depend on the spatial frequency of the superimposed noise, with high and low frequencies masking the text less effectively than intermediate frequencies (Solomon and Pelli 1994). Figure 1 illustrates the effect. The authors of the referred work related the frequency domain of most effective inhibition to the size of the letters: noise at three cycles per letter was found to be most effective.

The influence of the visual environment in which an object is embedded on our perception of the object is a central topic in psychophysics and perceptual psychology. It is a source of inspiration and questions for neuroscience and computational modelling. This certainly applies to the work referred to above. That study, however, left open two important questions. The first one is whether the observed effect is specific to text and which parameters determine the spatial frequency of most effective suppression. As demonstrated by Fig. 1, the effect seems not to be specific to text: next to letters, it can be observed simultaneously for other types of visual stimuli as well, such as object icons and bars. The second open question concerns the origin of the effect in terms of underlying neural processes. While the referred study contains an observation and a description of an interesting phenomenon in the tradition of psychophysics, a link to neurophysiology and an explanation in the style of computational neuroscience is still lacking. In the current paper we address the above questions.

Various visual illusions and psychophysical experiments have shown that the perception of an oriented stimulus, such as a line or a contour, can be influenced by the presence of other such stimuli (distractors) in

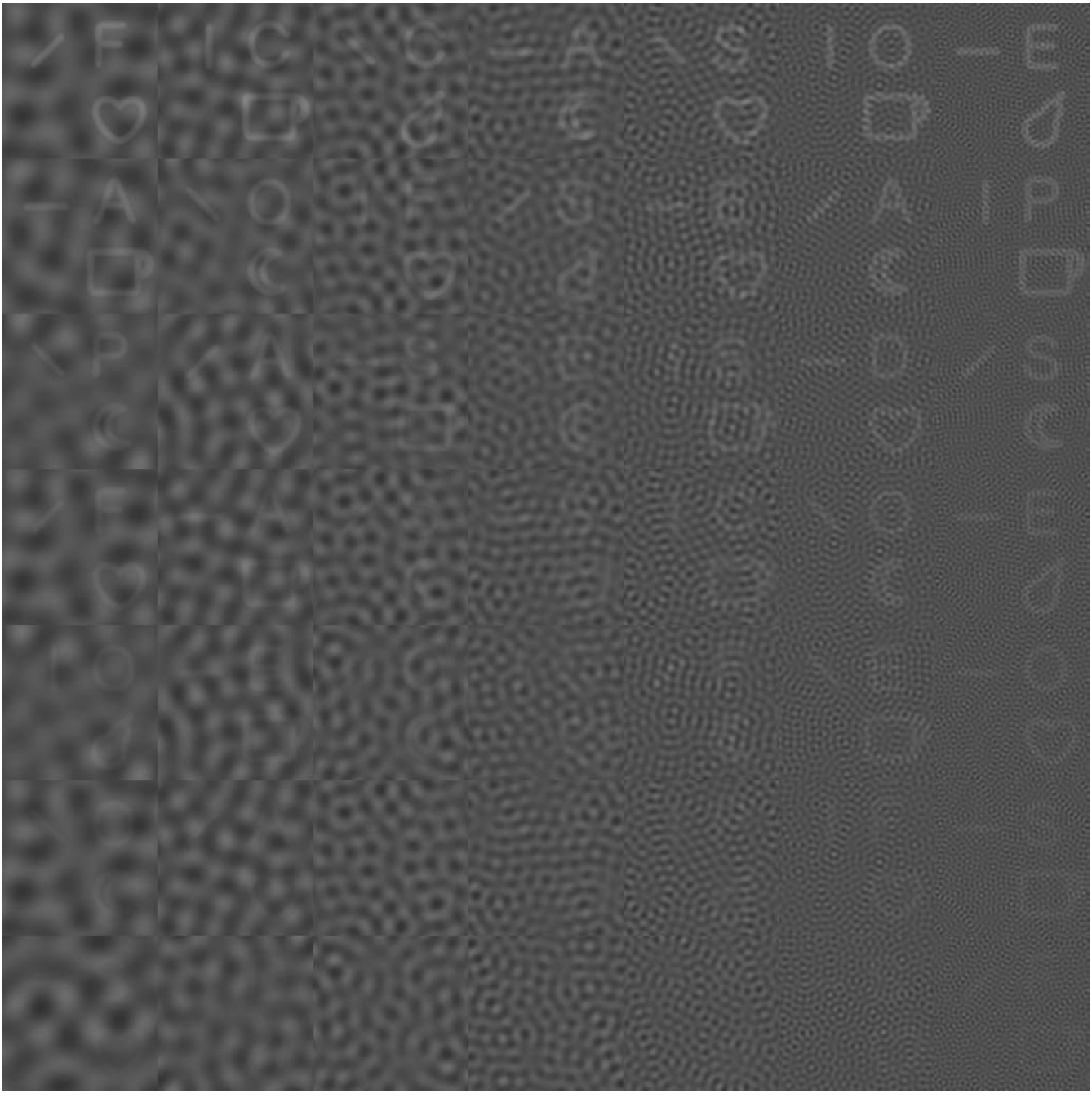

Fig. 1. The masking effect of band-spectrum noise on the perception of objects does not depend on the type of visual stimuli embedded in the noise: the same effect is observed simultaneously for letters, icons of familiar objects, and bars. On the *vertical axis* the contrast ranges from 0.05 to 0.25. On the *horizontal axis* the noise frequency ranges from 0.38 to 3.05 cycles per degree (cpd)

its neighborhood. This influence can, for instance, manifest itself in an overestimation of an acute angle between two lines (Blakemore et al. 1970) or in the so-called orientation contrast pop-out effect, Fig. 2a. See further, e.g., Nothdurft (1991). Figure 2b demonstrates how the saliency of a contour is reduced by a surrounding bar texture. Figure 2c illustrates another manifestation of the concerned influence – the suppression of contour perception by a grating. A part of a contour of a triangle ''disappears'' in the grating, and it does this most easily when the grating consists of elements similar to the contour with respect to orientation and width. The latter (width) aspect brings the spatial frequency of the grating into relation with the width of the contour. The spatial-frequency dependence of the text-masking effect illustrated by Fig. 1 might have a similar origin: parts of the stroke of a low-contrast letter can be hidden in a surrounding texture if the stroke width is comparable with the characteristic width of the light areas of the texture at places where the letter stroke and the texture are ''in phase.''

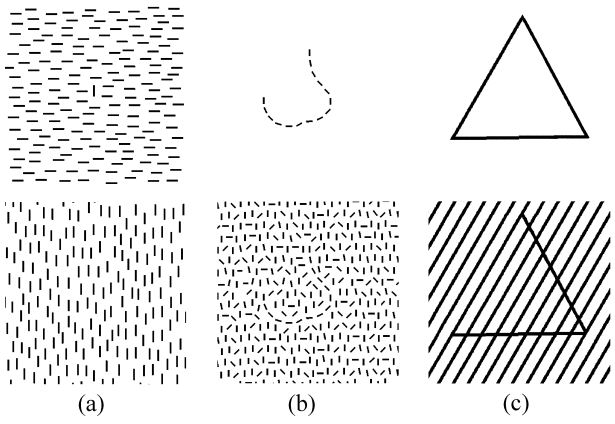

Fig. 2. Various manifestations of perception modulation by the context. **a** Orientation contrast pop-out: the oriented stimulus in the center segregates if it is surrounded by elements of orthogonal orientation. In contrast, it does not "pop out" in a surrounding of parallel stimuli. **b** A contour that is not surrounded by other stimuli is more salient than a contour that is surrounded by texture. **c** The three legs of the triangle are perceived quite differently, depending on the visual environment in which the triangle is embedded (Galli and Zama 1931). On the background of a grating of bars parallel to one of the legs of the triangle, that leg is less salient than the other two legs. The decreased saliency of a line embedded in a grating of parallel lines has been called "the social conformity of a line" by Kanizsa. For further examples, see (Kanizsa 1979)

Elsewhere, the perceptual effect illustrated by Fig. 2c was related to the function of a specific type of orientation-selective cell in the primary visual cortex for which the response to an optimal stimulus over the cell's classical (excitatory) receptive field (CRF) is influenced by other stimuli outside that area (Petkov and Kruizinga 1997). In contrast to other orientation-selective neurons, such as simple (Hubel and Wiesel 1968), complex (Hubel and Wiesel 1968), and grating (von der Heydt et al. 1992) cells, the concerned type of neuron responds strongly to single bars (of appropriate size and orientation) over its CRF and shows reduced response when a grating is added to the surround. The behavior of such a cell was systematically studied by Blakemore and Tobin (1972). The suppression effect they observed was strongest when the surround grating and the optimal bar stimulus over the CRF had the same orientation and decreased when the angle between the two was increased. This dependence on the orientation difference is well reproduced by a computational model of this type of cell proposed by Petkov and Kruizinga (1997).

Various neurophysiological studies have shown that response modulation by stimuli beyond the CRF is a common phenomenon in area V1 (Knierim and van Essen 1992; Li and Li 1994; Kapadia et al. 1995; Sillito et al. 1995; Zipser et al. 1996; Nothdurft et al. 1999). For instance, response suppression by a bar texture surround outside the CRF was found in nearly 80% of all orientation-selective cells reported in Knierim and van Essen (1992) and Nothdurft et al. (1999). Although only approximately one third of the cells showed dependence of the effect on the orientation of the texture surrounding the CRF, these authors suggested that this type of differential suppression might be the possible source of the pop-out effect in orientation contrast experiments (Fig. 2a). As illustrated by Fig. 2b, a surrounding texture need not have the same orientation as a central stimulus in order to suppress the saliency of a contour. This perceptual effect correlates well with the properties of 40% of the orientation-selective cells in area V1 for which the strength of non-CRF inhibition does not depend on the orientation of the surround stimuli (Nothdurft et al. 1999).

In this paper we report on results in two areas: psychophysics and computational modelling. As far as the psychophysical part of this study is concerned (Sect. 2), we had to redo and extend the experiments of Solomon and Pelli (1994), addressing two questions: (1) What does the spatial frequency domain of most effective masking depend on, the letter size (as suggested by Solomon and Pelli 1994) or the width of the letter stroke? (2) Is the effect specific to text or can it be observed for objects other than letters as well? As for the computational modelling part of the paper (Sect. 3), we introduce computational models of two types of orientation-selective cell with non-CRF inhibition and apply these models to the images used in the psychophysical experiments. We use the results to suggest a possible neural basis of the text-masking effect illustrated by Fig. 1 and the other perceptual effects shown in Fig. 2. Finally, in Sect. 4 we summarize and discuss the results.

2 Psychophysical experiments

2.1 Methods

We used images of letters with superimposed band-spectrum noise; see Fig. 3. A test image is generated as a

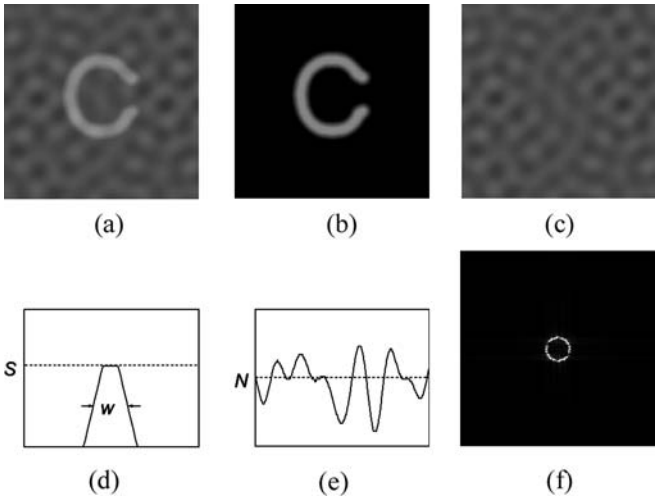

Fig. 3. A test image **a** is the sum of an image of a light letter on a black background **b** and a noise image **c**. The letter contour width w is measured at half height of the maximum luminance S of the trapezoid luminance profile of the stroke **d**. The noise image is a superposition of a constant component of luminance N and 100 sine waves of randomly selected different orientations and spatial frequencies within a narrow band. A characteristic 1-D luminance profile of the noise in the space domain is shown in **e**; its 2-D spectrum is presented in **f**

sum of an image of a light letter on a black background and a noise image (Fig. 3a–c). The letter contrast C is defined as the ratio $C = S/N$ of the maximum luminance S of the letter and the average luminance N of the noise. The letter contour width w is measured at half height of the trapezoid luminance profile of the contour (Fig. 3d). The noise image is a superposition of a constant component of luminance N and 100 sine waves (gratings) of randomly selected different orientations and spatial frequencies within a narrow band ($\Delta f = 0.14$ cycles per degree). The amplitude of the sine waves is chosen to be one-third of the average noise luminance N. The superposition of 100 such waves results in the noise-luminance profile illustrated in Fig. 3e. The two-dimensional (2-D) spectrum of the noise is shown in Fig. 3f.

Each test image had a size of 0.135×0.135 m on a computer screen (Toshiba Tecra 750CDT) and was viewed at a distance of 0.55 m, thus having visual angle dimensions of $14° \times 14°$. An observer was presented a series of such images of constant letter size and contour width and varying spatial frequency of the superimposed noise and luminance of the letter. Five letter contrasts ($C = 2^{k-6}$, $k = 0, \ldots, 4$) and eleven noise spatial frequencies ($f = 0.27 \cdot \sqrt{2}^k$ cycles per degree, $k = 0, \ldots, 10$) were used, resulting in 55 points in the spatial frequency-luminance contrast space, and 20 images were presented per point. Each image contained one of the letters A, C, D, E, F, O, P, R, S, or no letter (blank), all options represented with equal probability. The 1100 images generated in this way were presented in a random order to prevent accommodation to a specific spatial frequency, letter contrast or letter instance. Each image was presented for 1 s, and the observer had to indicate the letter seen. The recognition rate was determined as a two-variate function of the noise spatial frequency and the letter contrast for each of the 55 referred points. The curve in the spatial frequency-luminance contrast space, along which the recognition rate amounted to 90% (more precisely, between 77% and 100% with $p = 0.05$), was determined by interpolation. This experiment was carried out for different letter sizes ($h = 0.7°, 1.8°, 2.9°, 5.5°$) and contour widths ($w = 0.08°, 0.4°$). Instead of letters, the same set of experiments was carried out with object icons (heart, pear, half-moon, and cup) and with bars of different orientations: 0° (–), 45° (/), 90° (|) and 135° (\). All experiments were conducted with two observers. The results obtained for the two observers were consistent with each other within the statistical error.

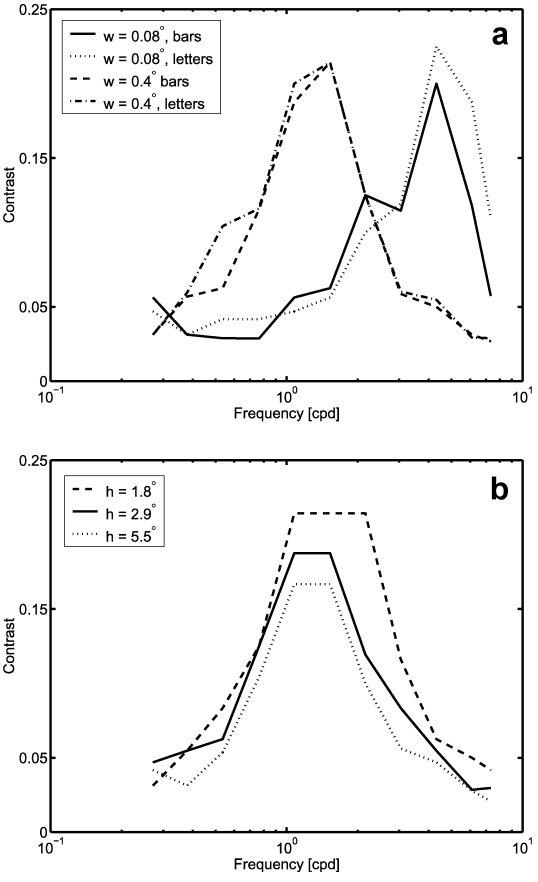

Fig. 4. **a** Bar and letter contrast at 90% recognition rate as a function of the spatial frequency of the noise. The stroke width is denoted by w (Observer 1). **b** Bar contrast at 90% recognition rate as a function of the spatial frequency of the noise; h denotes the bar length. Bar width was kept constant at $w = 0.4°$. The same results were obtained for letters and object icons (Observer 2)

2.2 Results

Figure 1 illustrates the simultaneousness of the effect for letters, bars, and object icons. It is obvious that the observed effect is not specific to the recognition of letters, which was the focus of the study by Solomon and Pelli (1994), but applies also to other objects and even to single bars. As illustrated quantitatively by Fig. 4a, the results for bars and letters are the same within the statistical error. The same results are obtained for object icons as well. These results suggest that an explanation of the observed effect should be sought in the low-level interaction of the band-limited noise with the contours of an object rather than with the object as a whole, be it a letter or not. Furthermore, as illustrated by Fig. 4a, the frequency domain of most effective inhibition depends on the stroke width of the object contour: the recognition of two bars or letters of the same size but different contour widths is most effectively inhibited in different frequency domains – the thinner the contours, the higher the characteristic domain of inhibition frequencies. In contrast, the frequency domain of most effective inhibition does not depend on the object size. The sole effect of object size on the inhibition is that of contrast rescaling: objects of a given size and contour width are correctly recognized at a lower contrast than the same objects of a smaller size and the same contour width. The object size has no essential effect on the general form of the curve that describes the dependence of the contrast on the spatial frequency of the noise (Fig. 4b).

3 Computational models and experiments

3.1 Simple and complex cells

3.1.1 Simple cells and Gabor filters. The spatial-summation properties of simple cells can be modelled by a family of 2-D Gabor functions (Daugman 1985). We use a modified parameterization according to Petkov and Kruizinga (1997) that takes into account restrictions found in experimental data. A receptive field function $g_{\lambda,\sigma,\theta,\varphi}(x,y), (x,y) \in \Omega \subset \mathbb{R}^2$, centered in the origin specifies the response to an impulse at point (x,y) and is defined as follows:

$$\begin{aligned} g_{\lambda,\sigma,\theta,\varphi}(x,y) &= \mathrm{e}^{-\frac{\tilde{x}^2+(\gamma\tilde{y})^2}{2\sigma^2}} \cos\left(2\pi\frac{\tilde{x}}{\lambda}+\varphi\right) \\ \tilde{x} &= x\cos\theta + y\sin\theta, \\ \tilde{y} &= -x\sin\theta + y\cos\theta \ , \end{aligned} \tag{1}$$

where $\gamma = 0.5$ is a constant, called the spatial aspect ratio, that determines the ellipticity of the receptive field. The standard deviation σ of the Gaussian factor determines the size of the receptive field. The parameter λ is the wavelength and $1/\lambda$ the spatial frequency of the cosine factor. They are called, respectively, the preferred wavelength and preferred spatial frequency of the Gabor function. The ratio σ/λ determines the spatial frequency bandwidth and the number of parallel excitatory and inhibitory stripe zones that can be observed in the receptive field (Fig. 5). In this paper, we fix the value of this ratio to $\sigma/\lambda = 0.56$, which corresponds to a bandwidth of one octave at half-response. The angle parameter $\theta, \theta \in [0,\pi)$, determines the preferred orientation. The parameter $\varphi, \varphi \in (-\pi,\pi]$, is a phase offset that determines the symmetry of $g_{\lambda,\sigma,\theta,\varphi}(x,y)$ with respect to the origin: for $\varphi = 0$ and $\varphi = \pi$ it is center symmetric (also called even), and for $\varphi = -\pi/2$ and $\varphi = \pi/2$ it is antisymmetric (or odd); all other cases are asymmetric mixtures.

The response $r_{\lambda,\sigma,\theta,\varphi}(x,y)$ of a simple cell with a receptive field function $g_{\lambda,\sigma,\theta,\varphi}(x,y)$ to an input image with luminance distribution $f(x,y)$ is computed by weighted integration:

$$\begin{aligned} &r_{\lambda,\sigma,\theta,\varphi}(x,y) \\ &\quad = \int\int_{\Omega} f(u,v) g_{\lambda,\sigma,\theta,\varphi}(u-x, v-y)\mathrm{d}u\mathrm{d}v \ . \end{aligned} \tag{2}$$

The model used in (Petkov and Kruizinga 1997) also involves half-wave rectification and contrast normalization, but we do not need these aspects of the function of simple cells in the context of this paper. In image processing and computer vision, the filter defined by Eqs. 1 and 2 is known as the (linear) Gabor filter.

As we shall see below, stimuli outside the CRF can have considerable influence on the response of a neuron. The model according to Eq. 2 does not take this effect into account and can be considered to be correct only for stimuli that are confined to the CRF. We define the CRF of a simple cell, which is centered at point (x,y), as the following elliptic region (Fig. 5):

$$\begin{aligned} \mathrm{CRF}_{\sigma,\theta,x,y} &= \{(u,v) \mid (\tilde{x}-\tilde{u})^2 + (\gamma(\tilde{y}-\tilde{v}))^2 \le 4\sigma^2\} \\ \tilde{x}-\tilde{u} &= (x-u)\cos\theta + (y-v)\sin\theta, \\ \tilde{y}-\tilde{v} &= -(x-u)\sin\theta + (y-v)\cos\theta \ . \end{aligned} \tag{3}$$

Outside this region the receptive field function $g_{\lambda,\sigma,\theta,\varphi}(u-x, v-y)$ of the concerned cell takes negligibly small values.

3.1.2 Complex cells and Gabor energy filters. The Gabor energy is related to a model of a complex cell that combines the responses of a pair of simple cells that have a phase difference of $\frac{\pi}{2}$. The results $r_{\lambda,\sigma,\theta,0}(x,y)$, and $r_{\lambda,\sigma,\theta,-\frac{\pi}{2}}(x,y)$ of a pair of a symmetric and an antisymmetric filter are combined in the Gabor energy $E_{\lambda,\sigma,\theta}(x,y)$ as follows:

$$E_{\lambda,\sigma,\theta}(x,y) = \sqrt{r^2_{\lambda,\sigma,\theta,0}(x,y) + r^2_{\lambda,\sigma,\theta,-\frac{\pi}{2}}(x,y)} \ . \tag{4}$$

In the following, we use the Gabor energies $E_{\lambda,\sigma,\theta_i}(x,y)$ for a number of N_θ different orientations

$$\theta_i = \frac{i\pi}{N_\theta}, \qquad i = 0,1,\ldots,N_\theta - 1 \ . \tag{5}$$

3.2 Models of non-CRF inhibition

Nearly 80% of all orientation-selective neurons in area V1 show some form of significant response modulation by texture surrounding their CRFs, with suppression being much more frequent than enhancement (Knierim and van Essen 1992; Nothdurft et al. 1999). On average over the whole population of such cells the response to an optimal stimulus over the CRF is reduced by 42% when texture is added to the surround.

We now extend the Gabor energy operator with an inhibition term to qualitatively reproduce the suppression influence that surround texture can have. The choice of this operator, which is a model of a complex cell, is merely for illustration purposes: the Gabor filter that is a model of a simple cell can be extended with a suppression term in a similar way.

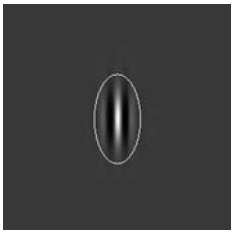

Fig. 5. Intensity map of a 2-D Gabor function that models the spatial-summation properties of a simple cell. *Gray levels*, which are lighter and darker than the background, indicate zones in which the function takes positive and negative values, respectively. The *bright ellipse* $x^2 + (\gamma y)^2 = 4\sigma^2$ specifies the boundary of the classical receptive field outside which the function takes negligibly small values

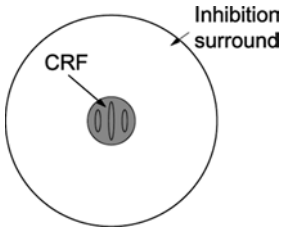

Fig. 6. Non-CRF inhibition is caused by the surround of the CRF, which is defined by the weighting function $w_\sigma(x,y)$

For a given point in the image, the inhibition term is computed by weighted summation of the responses in a ring-formed area surrounding the CRF centered at the concerned point (cf. Fig. 6). We use a normalized weighting function $w_\sigma(x,y)$ defined as follows:

$$w_\sigma(x,y) = \frac{1}{||H(\mathrm{DoG}_\sigma)||_1} H(\mathrm{DoG}_\sigma(x,y)),$$
$$H(z) = \begin{cases} 0 & z < 0 \\ z & z \geq 0 \end{cases}, \tag{6}$$

where $H(z)$ and $||.||_1$ denote half-wave rectification and the L_1 norm, respectively, and $\mathrm{DoG}_\sigma(x,y)$ is the following difference of Gaussian functions[1]:

$$\mathrm{DoG}_\sigma(x,y) = \frac{1}{2\pi(4\sigma)^2} \mathrm{e}^{-\frac{x^2+y^2}{2(4\sigma)^2}} - \frac{1}{2\pi\sigma^2} \mathrm{e}^{-\frac{x^2+y^2}{2\sigma^2}} . \tag{7}$$

We consider two types of inhibition: isotropic and anisotropic. In anisotropic inhibition, only neurons with the same preferred orientation as the concerned neuron contribute to the suppression. In isotropic inhibition, all surround neurons contribute to the suppression in an equal way, independently of their preferred orientations.

3.2.1 Anisotropic inhibition and bar cells. We model anisotropic inhibition by computing a suppression term $t_{\lambda,\sigma,\theta_i}(x,y)$ for each orientation θ_i as a convolution of the Gabor energy $E_{\lambda,\sigma,\theta_i}(x,y)$ for that orientation with the weighting function $w_\sigma(x,y)$:

$$t_{\lambda,\sigma,\theta_i}(x,y) = \int\int_\Omega E_{\lambda,\sigma,\theta_i}(u,v) \, w_\sigma(u-x, v-y) \, \mathrm{d}u\mathrm{d}v . \tag{8}$$

We now introduce a new operator $\tilde{b}^{A,\alpha}_{\lambda,\sigma,\theta_i}(x,y)$, which takes as its inputs the Gabor energy $E_{\lambda,\sigma,\theta_i}(x,y)$ and the inhibition term $t_{\lambda,\sigma,\theta_i}(x,y)$:

$$\tilde{b}^{A,\alpha}_{\lambda,\sigma,\theta_i}(x,y) = H(E_{\lambda,\sigma,\theta_i}(x,y) - \alpha t_{\lambda,\sigma,\theta_i}(x,y)) , \tag{9}$$

with $H(z)$ defined as in Eq. 6. The factor α controls the strength of the suppression exercised by the surround on the Gabor energy operator. Notice that one can model enhancement by taking negative values of α.

As to the function of this operator, one can say the following. If there is no texture in the surrounding of a given point, the response of this operator at that point will be similar to the response of the Gabor energy operator. (Actually, the response will be slightly reduced.) A line or edge passing through that point will be detected by the introduced operator in the same way as it is detected by the corresponding Gabor energy operator of appropriate orientation. However, if there are other lines and edges in the surroundings, the inhibition term $t_{\lambda,\sigma,\theta_i}(x,y)$ may become so strong that it can cancel completely the contribution of the Gabor energy operator, resulting in a zero response of the operator introduced above. The operator will thus respond to isolated lines and edges and show reduced or no response to textures of edges.

The computational model defined by Eq. 9 is similar to the bar cell operator introduced by Petkov and Kruizinga (1997)[2]. These authors introduced the term "bar cell" as a reference to the preference of the operator for bars vs. gratings. However, this term was not used to suggest a sharp functional division within the group of orientation-selective cells. Only in the extreme case of complete suppression will the operator respond exclusively to single bars and not respond to gratings of any periodicity. Such neurons have occasionally been encountered in neurophysiological studies (Schiller et al. 1976; von der Heydt et al. 1992). At the opposite end of the function continuum of orientation-selective cells, one can find the grating cells (von der Heydt et al. 1992), which respond optimally to gratings and, in the extreme case of "gratingness," do not respond to single bars.

Note that, for the type of suppression modelled by Eqs. 8 and 9, surround lines and edges of the same orientation as the main stimulus over the CRF will have a stronger suppression effect than such stimuli of other orientations, with suppression being weakest for surround stimuli that are orthogonal to the main stimulus. This type of differential response modulation for parallel vs. orthogonal surround texture has been observed by Nothdurft et al. (1999) in 24% of a large population of orientation-selective cells in area V1 of anesthetized macaque monkeys. These authors called such neurons "orientation contrast cells." In an additional 12% of the cells, the reverse effect of stronger suppression by an orthogonal surround texture was observed. The same authors called these latter neurons "uniform cells." Similar data were obtained for alert monkeys (Knierim and van Essen 1992): 32% orientation contrast cells and 6% uniform cells.

[1] The combination of standard deviations 4σ and σ was chosen in such a way that the zero crossings of $\mathrm{DoG}_\sigma(x,y)$ are near the circle $x^2+y^2=4\sigma^2$. More precisely, one can take $\tilde{x}^2+(\gamma\tilde{y})^2$ instead of x^2+y^2 as an argument on the right-hand side of Eq. 7 with $\tilde{x}$ and $\tilde{y}$ according to Eq. 1. The concerned area will then have an elliptic boundary like the one shown in Fig. 5. We did not do that for reasons of simplicity and computational efficiency and also because this did not influence the results substantially.

[2] The authors of that work use as an inhibition term the response of a grating cell, which is computed in a more intricate way from the simple or complex cell responses in the surround.

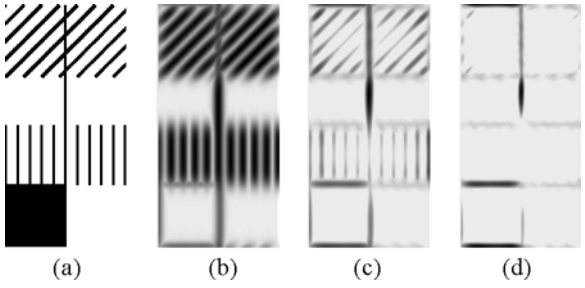

Fig. 7. a Synthetic input image. **b** The Gabor energy operator responds to lines and edges independently of the context, i.e., the surroundings in which these lines and edges are embedded (a max-value superposition of operators for different orientations is shown). **c,d** The anisotropic inhibition operator according to Eq. 10 responds differently to isolated lines and edges and lines that are surrounded by a grating of a different orientation on the one hand and bar gratings on the other. It leaves relatively unaffected the former stimuli and suppresses the latter ones. As a result, the responses to one and the same line – consider the vertical line in the middle of the input image – can be different, depending on the surroundings, the context, in which the line is embedded. The response to a line surrounded by similar parallel lines is reduced **c**, which might be the reason for the decreased perceptual saliency of such lines. In the extreme case, corresponding to large values of the inhibition coefficient α, the operator does not respond at all to gratings and lines that comprise a part of such gratings **d**. The response reduction observed in neurophysiological measurements can vary from moderate to very strong. For instance, it can amount to 90% of the response to an optimal stimulus over the CRF (Nothdurft et al. 1999)

The various terms mentioned above – bar cell, orientation contrast cell, and uniform cell – emphasize different aspects of the function of an orientation-selective neuron with anisotropic non-CRF inhibition. (At the same time, they illustrate the risks and limitations of naming a neuron type after a stimulus: such naming focuses attention on only one aspect of the neuron's function.)

For visualization purposes we now construct a new map $b^{A,\alpha}_{\lambda,\sigma}(x,y)$ with values of maximum response of $\tilde{b}^{A,\alpha}_{\lambda,\sigma,\theta_i}(x,y)$ over all orientations:

$$b^{A,\alpha}_{\lambda,\sigma}(x,y) = \max\{\tilde{b}^{A,\alpha}_{\lambda,\sigma,\theta_i}(x,y) \mid i = 0,1,\ldots,N_\theta - 1\} \ , \tag{10}$$

Defined in this way, the concerned operator will respond to isolated lines, edges, and bars of any orientation, but it will give a reduced response or not respond at all to groups of such stimuli that make part of a grating. See Fig. 7c,d.

Figure 8 shows the output of the above-defined operator for the input images of Fig. 2c and different (positive) values of the coefficient α that controls the strength of suppression. This figure illustrates how the anisotropic inhibition of the type modelled by Eqs. 8 and 9 can account for the effect of decreased saliency of contours embedded in a surrounding of lines of the same orientation. In the same way, this type of suppression can account for the orientation contrast pop-out effect (Fig. 9). Of course, this holds for the so-called "orientation contrast cells" or, equivalently, "bar cells". For "uniform cells" the opposite effect should be expected[3]. Due to the higher incidence of "orientation contrast cells" vs. "uniform cells" the overall effect over the whole population of anisotropic inhibition cells should be similar to the one illustrated by Fig. 8 and Fig. 9.

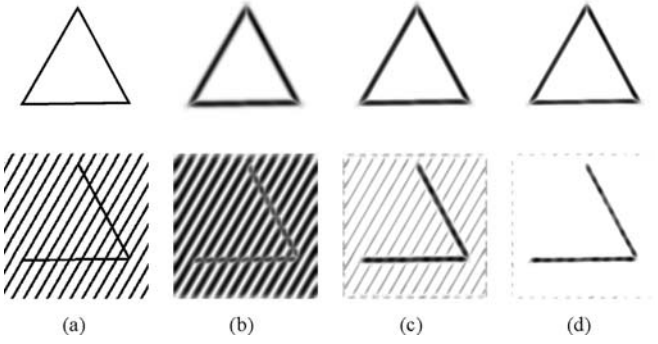

Fig. 8. Input images **a** and outputs of the anisotropic inhibition operator for different strengths of suppression: **b** $\alpha = 0$ (no suppression), in this case the operator is equivalent to a superposition of complex cell operators; **c** $\alpha = 2$; **d** $\alpha = 3.5$. The saliency of the leg of the triangle in the **bottom row**, which is parallel to the surrounding line grating, decreases with increasing strength of the suppression

3.2.2 Isotropic inhibition and general suppression cells. Nothdurft et al. (1999) found no significant difference in the modulation of the response to a center line stimulus by surround textures that were parallel or orthogonal to the center stimulus in 40% of the V1 cells they studied in monkeys under anesthesia. (In a previous study by Knierim and van Essen (1992), 27% of such cells were found in alert monkeys.) They found that suppression was far more frequent than enhancement. We refer to this type of suppression as isotropic inhibition[4] and model it by taking as an inhibition term the sum of the inhibition terms according to Eq. 8 over all possible orientations:

$$\tilde{b}^{I,\beta}_{\lambda,\sigma,\theta_i}(x,y) = H\Big(E_{\lambda,\sigma,\theta_i}(x,y) - \frac{\beta}{N_\theta}\sum_{i=0}^{N_\theta-1} t_{\lambda,\sigma,\theta_i}(x,y)\Big) \ . \tag{11}$$

The factor β controls the strength of the suppression exercised by the surround on the Gabor energy operator.

For visualization purposes we now construct a new map $b^{I,\beta}_{\lambda,\sigma}(x,y)$ with values of maximum response of $\tilde{b}^{I,\beta}_{\lambda,\sigma,\theta_i}(x,y)$ over all orientations:

$$b^{I,\beta}_{\lambda,\sigma}(x,y) = \max\{\tilde{b}^{I,\beta}_{\lambda,\sigma,\theta_i}(x,y) \mid i = 0,1,\ldots,N_\theta - 1\} \ . \tag{12}$$

This operator responds to isolated lines, edges, and bars of any orientation, but it gives a reduced response or

[3] In this case, the inhibition term $t_{\lambda,\sigma,\theta_i}(x,y)$ in Eq. 8 is computed by integrating $E_{\lambda,\sigma,\theta_i+\pi/2}(u,v)$ instead of $E_{\lambda,\sigma,\theta_i}(u,v)$.

[4] The term used by Nothdurft et al. (1999) to refer to the neurons that exhibit this type of inhibition is "general suppression cells."

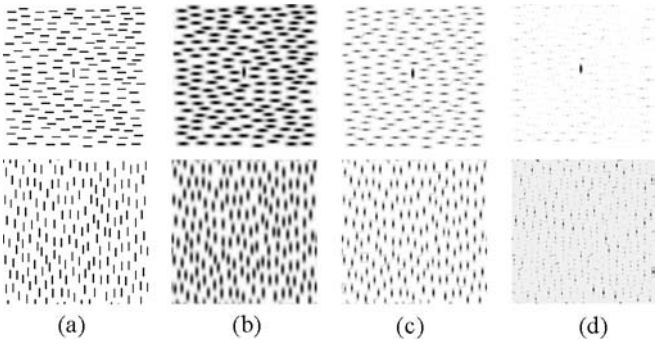

Fig. 9. Input images **a** and outputs of the anisotropic inhibition operator for different strengths of suppression: **b** $\alpha = 0$, **c** $\alpha = 2$, **d** $\alpha = 3.5$. The larger the suppression, the stronger the pop-out effect

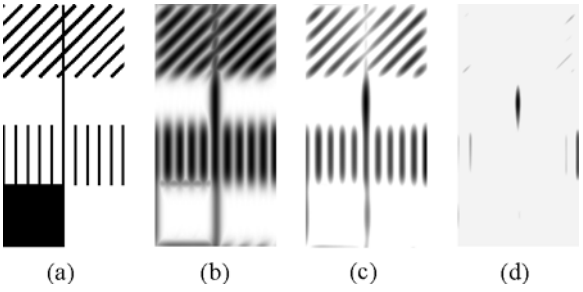

Fig. 10. **a** Synthetic input image. **b** The Gabor energy operator responds to lines and edges independently of the surroundings in which these lines and edges are embedded. **c,d** The isotropic inhibition operator leaves relatively unaffected isolated lines and edges but suppresses gratings and lines surrounded by gratings. Depending on the strength of the inhibition coefficient β, the response to a line embedded in texture can be reduced (**c**: $\beta = 7$) or completely suppressed (**d**: $\beta = 13.5$). In neurophysiological measurements, the strength of this type of suppression was found to vary considerably from neuron to neuron, from mild to strong and even complete suppression (Nothdurft et al. 1999)

does not respond at all to such stimuli if they are embedded in a texture of oriented stimuli, independent of the orientation of these latter stimuli (Fig. 10c,d). While isotropic suppression cannot account for orientation contrast pop-out, it may explain the decreased saliency of contours embedded in texture (Fig. 11).

3.3 Results

We applied the computational models defined by Eqs. 9 and 11 to the images we used in the psychophysical experiments with the human observers. We used letters and presented combinations of six letter contrasts and eleven noise spatial frequencies, resulting in 66 points in the spatial frequency-letter contrast space. Per point, ten images were used as input to the computational models.

Figure 12 shows test images (in the first and third rows) and the corresponding responses (in the second and fourth rows) of the anisotropic inhibition operator according to Eq. 10 for two spatial frequencies of the noise and six different letter contrasts. The test images contain the letter A with contour stroke width

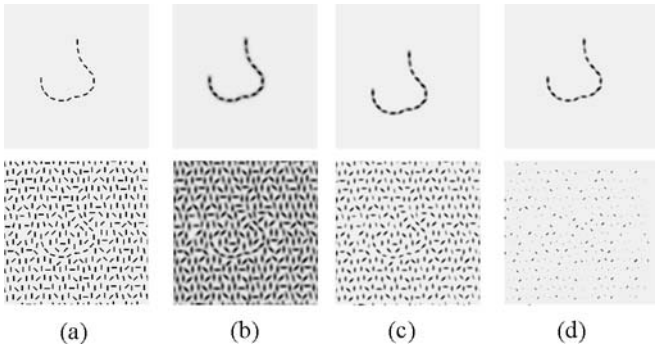

Fig. 11. Input images **a** and outputs of the isotropic inhibition operator for different strengths of suppression: **b** $\beta = 0$ (no inhibition), **c** $\beta = 3$, **d** $\beta = 6$. The decreased saliency of the contour embedded in texture (*bottom-left image*) can be explained by the decreased response of the corresponding cells to the contour segments in the presence of texture in their surroundings

$w = 0.08°$ and background noise with spatial frequencies of 4.3 cpd for the first row and 0.5 cpd for the third row. The contrast varies from $C = 0.5$ for the leftmost image down to $C = 0.0156$ for the rightmost image in a row.

In general, the operator removes texture noise while retaining letter contours; see the operator output in the last row of images in Fig. 12. However, at low contrasts and certain frequencies, which are related to the stroke width of the letter contour, the noise cannot be filtered without removing the letter contours as well (see second row of images in that figure). At such frequencies, letter contours can pass the filter only if they have sufficiently large contrast, as, for instance, in the leftmost image of the concerned row. As can be seen, the noise with spatial frequency of 4.3 cpd (in the first row in Fig. 12) results in higher suppression, and the letter is unrecognizable in the output image already for a relatively high value of the contrast $C = 0.25$. Noise with spatial frequency 0.5 cpd (see the third row in Fig. 12) does not obstruct the detection of the letter contour by the operator that much. Only when the contrast falls to $C = 0.0156$ does the response of the operator reduce to almost zero.

Based on the response images of the operator, we determined the curve in the spatial frequency-letter contrast space along which the recognition rate amounted to 90%. Correct recognition in this case means that the letter contour remains connected in the output of the operator so that it can be recognized by a human viewer. This is not always the case: the contour may be obscured in the output by responses of the operator to the background noise. For instance, the letter is not recognizable in the operator output shown in the second image in the second row of Fig. 12 because the contour is too disconnected.

The utility of the operator output for recognition is thus done subjectively by a human observer. The recognition decision can, however, easily be automated, i.e., it can be taken by a computer program, e.g., by first thresholding and binarizing an output image in such a way that the number of pixels assigned the value 1 in the

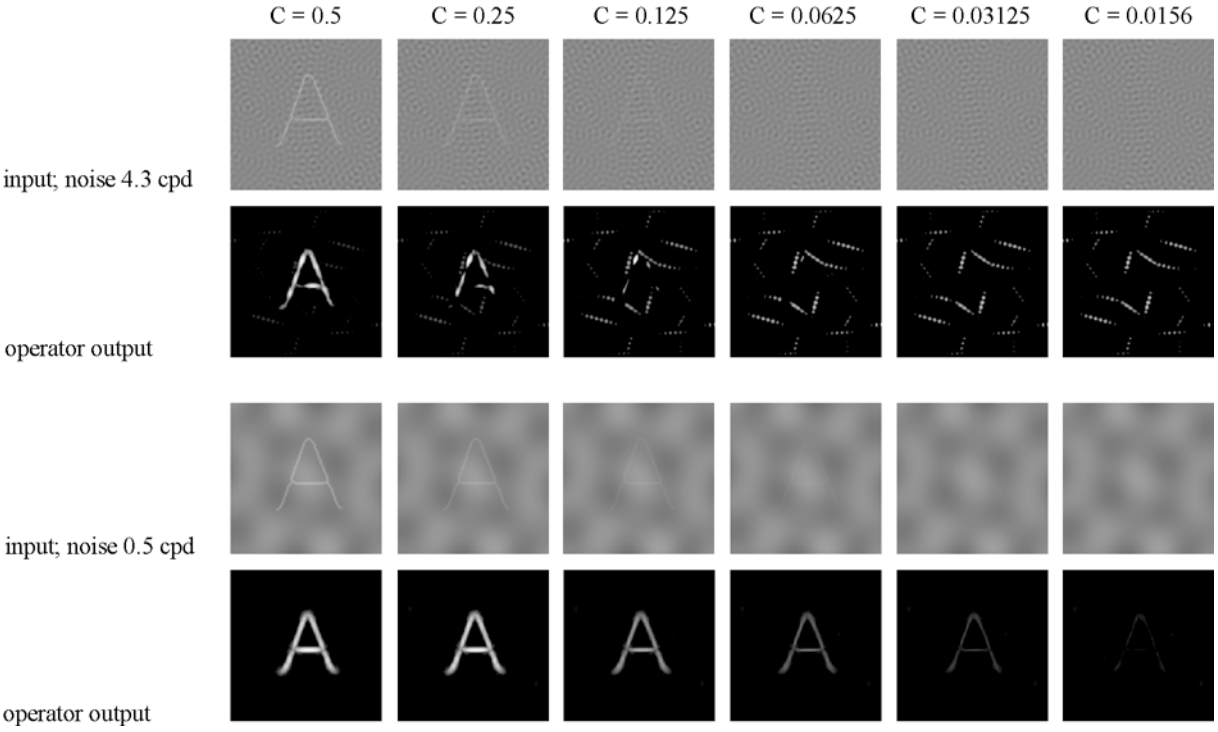

Fig. 12. Input stimuli (*odd rows*) and corresponding responses of the anisotropic inhibition operator (*even rows*) computed according to Eq. 10 with a suppression factor $\alpha = 2.0$. All input images contain the letter A with a stroke width $w = 0.08°$. Six letter contrasts were used, left to right, $C = 0.5, 0.25, 0.125, 0.0625, 0.03125, 0.0156$. The input images shown in the *first row* contain noise with a spatial frequency of 4.3 cpd. The noise cannot effectively be removed by the operator without removing the letter contour as well, and already for $C = 0.25$ the letter is unrecognizable in the output (*second row*). The *third row* shows input images with superimposed noise with a spatial frequency of 0.5 cpd, and the *fourth row* shows the corresponding responses. The operator effectively removes the background noise while letting the letter contour pass, except for the *bottom-right image* in which the operator response reduces to almost zero. Similar results were obtained with the isotropic inhibition model according to Eq. 11

resulting binary image is equal to the number of pixels occupied by a character in an input image. If a sufficiently large number of 1-pixels is found in the area of the image defined by a binary template of a given character, that character is considered as correctly recognized. This will be the case for images of high letter contrast. For low letter contrasts, most of the 1-pixels in the binarized output image will be due to noise, and the template-matching test will give a negative result. In this way, the usefulness of the operator output for recognition can be evaluated objectively.

The results of the operator output evaluation for letters of two contour widths are shown in Fig. 13. For each letter contour width there is a range of spatial frequencies of the noise for which, unless the letter contrast is high enough, the operator cannot effectively suppress the noise without suppressing the letter contour as well. The performance of the operator correlates astonishingly well with the perceptual sensitivity of humans measured in the psychophysical experiments (compare Fig. 13 and Fig. 4a). Similar results are obtained with an operator for which the inhibition by orthogonal texture is stronger than the suppression caused by parallel texture, as found in 12% of the cells studied by Nothdurft et al. (1999), and also with the isotropic inhibition model. This reveals non-CRF inhibition – of both types, isotropic and anisotropic – as the neural basis of the perceptual effect illustrated by Fig. 1.

4 Summary and discussion

We conducted a series of psychophysical experiments that demonstrate that the frequency domain of most effective inhibition of the visual perception of printed letters superimposed on a background of band-limited noise depends on the stroke width of the letter contour rather than on the letter size, as suggested previously by Solomon and Pelli (1994). The sole effect of object size on the inhibition turned out to be that of contrast rescaling: objects of a given size and contour width are correctly recognized at a lower contrast than the same objects of a smaller size and the same contour width. However, the object size has no essential effect on the form of the curve that describes the contrast dependence on the spatial frequency of the noise. We also demonstrated that the observed effect is not specific to the recognition of letters but also applies to other objects and even to single bars.

From these results we conclude that the observed effect need not be attributed specifically to a particular higher cognitive ability, such as the ability to recognize letters or to read. Furthermore, the fact that the same spatial-frequency dependence is observed not only for visual objects defined by contours but also for simple primitives, such as bars, means that this dependence is due to the properties of the visual system at a relatively low level at which bars, lines, and contours are detected (presumably

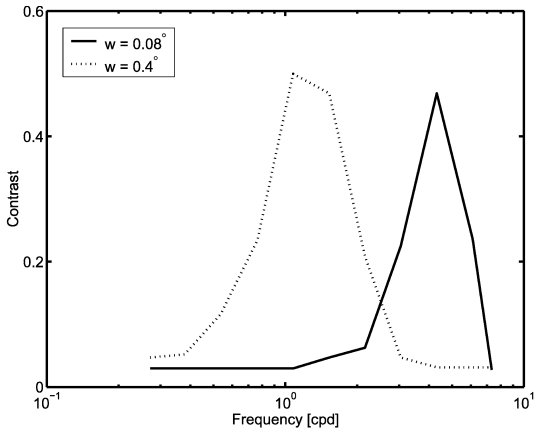

Fig. 13. Letter contrast as a function of the spatial frequency of the noise. The plot shows letter contrasts for which the anisotropic inhibition operator filters out noise and allows a sufficient part of the letter contour to pass so that the letter can be recognized (automatically or by a human observer) in the output of the operator in 90% of the cases. The performance of the operator correlates almost perfectly with the psychometric curves shown in Fig. 4a. Similar results were obtained with the isotropic inhibition model

V1) rather than to its properties at higher levels where integration of information and recognition takes place. The dependence of the detection contrast on the object size can be explained in probabilistic terms: at low contour contrast, some segments of the contour "disappear" in the noise; this occurs at places where contour and noise are "in phase." Other contour segments are, however, still visible. The bigger the object, the larger the number of visible contour segments and the higher the probability to integrate them into a percept of the object.

Furthermore, we related the observed effect to certain neural mechanisms and properties of neurons in visual area V1. More specifically, we attribute this effect to non-CRF inhibition. This mechanism has previously been suggested as the possible origin of other perceptual effects, such as the orientation contrast pop-out (Fig. 2a) (Knierim and van Essen 1992; Nothdurft et al. 1999) and the suppression of modal contours (Fig. 2c) (Petkov and Kruizinga 1997). The visual contexts involved in these perceptual effects can be considered Gestalt grouping contexts. Since, on the one hand, Gestalt grouping processes are traditionally thought to reside at late levels of cortical processing and we, on the other hand, demonstrate how these effects can be explained by cortical mechanisms residing at fairly early levels, we think that the concerned mechanisms of non-CRF inhibition provide an interesting basis for possible coupling of low-level mechanisms and high-level processes.

Our approach is computational: we introduced computational models of two types of cell that incorporate non-CRF inhibition. The models are based on Gabor energy filters - models of simple and complex cells – that are extended by surround inhibition of two types: anisotropic and isotropic. Anisotropic inhibition originates from surround neurons with the same preferred orientation as a neuron under consideration. This type of inhibition causes decreased responses to modal contours, i.e., contours embedded in texture of oriented stimuli of the same orientation as the contour. Isotropic inhibition originates from all orientation-selective neurons in the surroundings of a given neuron, and it does not depend on the preferred orientations of these neurons. This type of inhibition leads to decreased responses to both modal and amodal contours, i.e., to contours embedded in texture of any orientation.

We applied these computational models to the images we used in the psychophysical experiments. The computational results show a dependence of the inhibition effect on the spatial frequency of the background noise that is similar to the inhibition effect measured in the psychophysical experiments. This dependence was the same for the anisotropic and isotropic inhibition model. In this way, while the perceptual effects that involve Gestalt grouping contexts based on orientation similarity (see Fig. 2a and c) can be explained only by anisotropic non-CRF inhibition, the text-masking effect (Fig. 1) and the contour suppression effect illustrated by Fig. 2b can be explained by the computational models of both isotropic and anisotropic non-CRF inhibition.

The primary purpose of computational modelling is to provide a common quantitative explanation of a number of related observations and a means to predict the behavior of a system in new situations. An interesting question is whether one can interpret a model in order to gain an insight into the origins of the model. Newton's law of gravitation, for instance, has been very successful ever since it was proposed, despite the fact that the bare formula itself does not offer any ground for interpretation or any insight into the possible underlying source of gravitation. In this respect, one must be very careful with possible interpretations of the equations Eqs. 9 and 11 that embody the proposed computational models. A straightforward interpretation might be that the response of a V1 neuron is suppressed by the activity of other V1 neurons by inhibitory interconnections. Long horizontal connections in the striate cortex can connect neurons at distances by far exceeding their classical receptive fields (Gilbert and Wiesel 1989) and were shown to link cells with similar preferred orientations (Ts'o et al. 1986; Ts'o and Gilbert 1988). When such connections end on inhibitory interneurons (McGuire et al. 1991), they can produce the suppression observed in neurophysiological experiments (Knierim and van Essen 1992; Nothdurft et al. 1999) and modelled in the equations proposed in this paper. However, as shown by Lamme et al. (1997), feedback from higher cortical areas may also play a role in non-CRF inhibition. The latency differences between isotropic and anisotropic suppression found by Nothdurft et al. (1999) can also be considered as evidence for the involvement of feedback, at least as far as anisotropic inhibition is concerned. For further discussion on the role of long horizontal vs. feedback connections, we refer to Nothdurft et al. (1999).

There is one particular aspect of the computational models according to Eqs. 9 and 11 that requires special comment: the inhibition of a neuron with preferred spatial frequency $1/\lambda$ is caused by neurons with the same

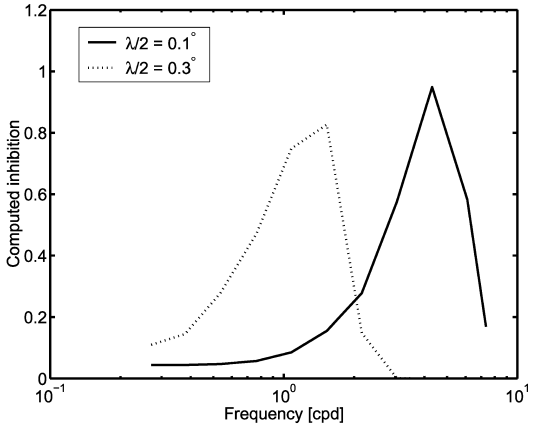

Fig. 14. Spatial frequency tuning curves of two Gabor functions with different preferred spatial frequencies. The preferred spatial frequency of a Gabor function is selected in such a way that the width ($\lambda/2$) of the central excitatory lobe of the function is (approximately) equal to the width of the contours used in the psychophysical experiments. The match of these curves with the psychometric curves (Fig. 4a) provides further evidence for the plausibility of the computational models according to Eqs. 9 and 11

preferred frequency. One can verify the plausibility of this assumption in the following way. If we assume that a contour of given width w is detected primarily by the activity it causes in neurons with a certain preferred spatial frequency $1/\lambda$ that is related to w, then the strength of the inhibition by superimposed band-limited noise will follow the spatial frequency tuning curve of the inhibiting neurons. The responses of these neurons are modelled by Gabor functions according to Eq. 1. Figure 14 shows the spatial frequency tuning curves of two such functions with different preferred spatial frequencies. The good match of these curves with the psychometric curves that specify the actual suppression effect that was measured in the psychophysical experiments (Fig. 4a) provides further evidence for the plausibility of the proposed models.

Finally, it was suggested by Nothdurft et al. (1999) that non-CRF inhibition may play a certain role in texture segmentation by generating sensitivity to (texture) feature contrast. Our computational model supports this assumption. The anisotropic surround inhibition operator (Eq. 9), for instance, will respond more strongly at the boundary between two textures of different preferred orientations than inside uniform texture regions. Both the isotropic (Eq. 11) and the anisotropic (Eq. 9) operator will respond more strongly at the boundary between two textures of different spatial frequency spectra than inside these textures. Besides this role of non-CRF inhibition, we think that the biological utility of non-CRF inhibition is the separation of contour from texture information and the mediation of object contours to higher cortical areas, as previously proposed by Petkov and Kruizinga (1997, 2000). Indeed, while non-CRF inhibition scarcely influences the responses to isolated contours, i.e., contours that are not embedded in texture, it strongly reduces the responses to texture. The biologically (in our view) most important effect of this neural mechanism might thus be that of contour (vs. texture) detection.

References

Blakemore C, Carpenter RHS, Georgeson MA (1970) Lateral inhibition between orientation detectors in the human visual system. Nature 228: 37–39

Blakemore C, Tobin EA (1972) Lateral inhibition between orientation detectors in the cat's visual cortex. Exp Brain Res 15: 439–440

Daugman JG (1985) Uncertainty relations for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J Opt Soc Am A 2: 1160–1169

Galli A, Zama A (1931) Untersuchungen über die Wahrnehmung ebener geometrischer Figuren, die ganz oder teilweise von and eren geometrischen Figuren verdeckt sind, zeitschrift für Psychologie 123: 308–348

Gilbert CD, Wiesel TN (1989) Columnar specifity of intrinsic horizontal and corticocortical connections in cat visual cortex. J Neurosci 9: 2432–2442

Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. J Physiol Lond 195: 215–243

Kanizsa G (1979) Organization in Vision, Essays on Gestalt Perception, Praeger, New York

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

Knierim JJ, van Essen DC (1992) Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J Neurophysiol 67: 961–980

Lamme VAF, Zipser J, Spekreijse H (1997) Figure-ground signals in V1 depend on extrastriate feedback. Invest Ophthalmol Vis Sci 38: S969

Li CY, Li W (1994) Extensive integration field beyond the classical receptive field of cat's striate cortical neurons – classification and tuning properties. Vis Res 34: 2337–2355

McGuire BA, Gilbert CD, Rivlin PK, Wiesel TN (1991) Targets of horizontal connections in macaque primary visual cortex. J Comp Neurol 305: 370–392

Nothdurft HC (1991) Texture segmentation and pop-out from orientation contrast. Vis Res 31: 1073–1078

Nothdurft HC, Gallant JL, van Essen DC (1999) Response modulation by texture surround in primate area V1: correlates of "popout" under anasthesia. Vis Neurosci 16: 15–34

Petkov N, Kruizinga P (1997) Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells. Biol Cybern 76(2): 83–96

Petkov N, Kruizinga P (2000) Perception of form and texture through complementary bar and grating cell channels. Perception 29(suppl): 60c

Schiller PH, Finlay BL, Volman SF (1976) Quantitative studies of single-cell properties in monkey striate cortex. III. Spatial frequencies. J Neurophysiol 39: 1334–1351

Sillito AM, Grieve KL, Jones HE, Gudeiro J, Davis J (1995) Visual cortical mechanisms detecting focal orientation discontinuities. Nature 378: 492–496

Solomon JA, Pelli DG (1994) The visual filter mediating letter identification. Nature 369: 395–397

Ts'o DY, Gilbert CD (1988) The organization of chromatic and spatial interactions in the primate striate cortex. J Neurosci 8: 1712–1727

Ts'o DY, Gilbert CD, Wiesel TN (1986) Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. J Neurosci 6: 1160–1170

Von der Heydt R, Peterhans E, Dürsteler MR (1992) Periodic-pattern-selective cells in monkey visual cortex. J Neurosci 12: 1416–1434

Zipser J, Lamme VAF, Schiller PH (1996) Contextual modulation in primary visual cortex. J Neurosci 16: 7376–7389