

Improved Contour Detection by Non-Classical Receptive Field Inhibition

Cosmin Grigorescu, Nicolai Petkov, and Michel A. Westenberg

Institute of Mathematics and Computing Science
University of Groningen,
P.O. Box 800, 9700 AV Groningen, The Netherlands.
cosmin@cs.rug.nl, petkov@cs.rug.nl, michel@cs.rug.nl

Abstract. We propose a biologically motivated computational step, called non-classical receptive field (non-CRF) inhibition, to improve the performance of contour detectors. We introduce a Gabor energy operator augmented with non-CRF inhibition, which we call the bar cell operator. We use natural images with associated ground truth edge maps to assess the performance of the proposed operator regarding the detection of object contours while suppressing texture edges. The bar cell operator consistently outperforms the Canny edge detector.

1 Introduction

In the early 1960s, an important finding in the neurophysiology of the visual system of monkeys and cats was that the majority of neurons in the primary visual cortex function as edge detectors. Such neurons react strongly to an edge or a line of a given orientation in a given position of the visual field [1]. The computational models of two types of orientation selective cell, called the simple cell and the complex cell, which were developed, gave the basis for biologically motivated edge detection algorithms in image processing. In particular, a family of two-dimensional Gabor functions was proposed as a model of the receptive fields of simple cells [2] and subsequently used widely in various image processing tasks, such as image coding and compression, face recognition, texture analysis, and edge detection.

The behaviour of orientation selective cells has turned out to be more complex than suggested by early measurements and models. In particular, the concept of a receptive field — the region of the visual field in which an optimal stimulus elicits response from a neuron — had to be reconsidered. This region is presently referred to as the classical receptive field (CRF). Detailed studies have shown that once a cell is activated by a stimulus in its CRF, another, simultaneously presented stimulus outside that field can have effect on the cell response (cf. Fig. 1(a)). This, mostly inhibitive effect is referred to as *non-classical receptive field inhibition*, and it is exhibited by 80% of the orientation selective cells [3]. In general, an orientation selective cell with non-CRF inhibition responds most strongly to a single bar, line, or edge in its receptive field, and shows reduced response when other such stimuli are present in the surrounding. In an extreme case, the cell responds only to an isolated bar or line. Such cells have

H.H. Bülthoff et al. (Eds.):BMCV 2002, LNCS 2525, pp. 50-59, 2002.

©Springer-Verlag Berlin Heidelberg 2002

been found by neurophysiologists: Schiller et al. [4] found many cells in area V1 which responded strongly to single bars and edges but did not respond to sine-wave gratings. Similar cells were encountered by Peterhans and Von der Heydt [5]. This type of cell was called the *bar cell* and a computational model was proposed for it elsewhere [6].

The above mentioned neurophysiological behaviour of bar cells correlates well with the results of various psychophysical experiments, which have shown that the perception of an oriented stimulus, such as a line, can be influenced by the presence of other such stimuli (distractors) in its neighbourhood. This influence can, for instance, manifest itself in an overestimation of an acute angle between two lines [7], or in an orientation pop-out effect, Fig. 1(b), or in a decreased saliency of groups of parallel lines [8]. Figure 1(c) illustrates the latter effect, where the perception of a contour is suppressed by a grating.

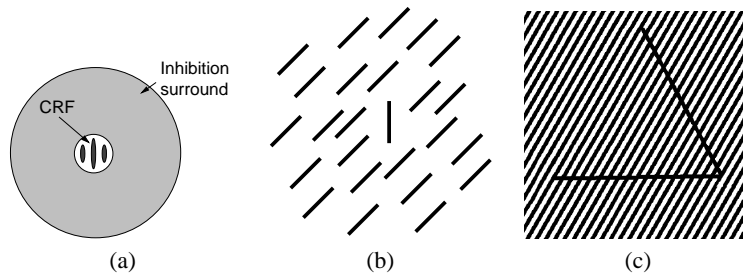


Fig. 1. (a) Non-CRF inhibition is caused by the surround of the CRF. (b) The pop-out effect of an oriented line segment on a background of other segments (distractors): the segment pops out only if its orientation is sufficiently different from that of the background. (c) The three legs of the triangle are not perceived in the same way: the leg which is parallel to the bars of the grating does not pop out as the other two legs.

In this paper, we examine the role of the non-CRF inhibition mechanism in the process of edge detection and its potential usefulness in image processing and computer vision. Our main hypothesis is that this mechanism suppresses edges which make part of texture, while it does not suppress edges that belong to the contours of objects. An edge detector which employs this inhibition mechanism will thus be more useful for contour-based object recognition tasks, such as shape comparison [9], than traditional edge detectors, which do not distinguish between contour and texture edges.

The paper is organized as follows. Section 2 describes the computational model. The simple cell and complex cell models and the related Gabor and Gabor energy filters are briefly discussed, and the bar cell operator is introduced. In Section 3, we evaluate the performance of the bar cell operator, and compare it to the Canny edge detector. Finally, we discuss possible extensions of the proposed model in Section 4.

2 Computational Model

2.1 Simple Cells and Gabor Filters

The spatial summation properties of simple cells can be modeled by a family of two-dimensional Gabor functions [2]. We use a modified parameterization to take into account restrictions found in experimental data [6]. A receptive field function of such a cell, in engineering terms the impulse response, $g_{\lambda,\sigma,\theta,\varphi}(x, y)$, $(x, y) \in \Omega \subset \mathbb{R}^2$, which is centered around the origin, is given by:

$$g_{\lambda,\sigma,\theta,\varphi}(x, y) = e^{-\frac{\tilde{x}^2 + \gamma^2 \tilde{y}^2}{2\sigma^2}} \cos\left(2\pi \frac{\tilde{x}}{\lambda} + \varphi\right) \quad (1)$$

$$\tilde{x} = x \cos \theta + y \sin \theta, \quad \tilde{y} = -x \sin \theta + y \cos \theta,$$

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 linear size of the receptive field. The parameter λ is the wavelength and $1/\lambda$ the spatial frequency of the cosine factor. The ratio σ/λ determines the spatial frequency bandwidth, and, therefore, the number of parallel excitatory and inhibitory stripe zones which can be observed in the receptive field. In this paper, we fix the value of the ratio σ/λ to $\sigma/\lambda = 0.56$, which corresponds to a half-response bandwidth of one octave. The angle parameter θ , $\theta \in [0, \pi)$, determines the preferred orientation. The parameter φ , $\varphi \in (-\pi, \pi]$, 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 symmetric (or even), and for $\varphi = -\frac{\pi}{2}$ and $\varphi = \frac{\pi}{2}$ it is antisymmetric (or odd); all other cases are asymmetric mixtures.

The response $r_{\lambda,\sigma,\theta,\varphi}(x, y)$ of a simple cell at position (x, y) to an input image with luminance distribution $f(x, y)$ is computed by convolution:

$$\begin{aligned} r_{\lambda,\sigma,\theta,\varphi}(x, y) &= f(x, y) * g_{\lambda,\sigma,\theta,\varphi}(x, y) \\ &= \iint_{\Omega} f(u, v) g_{\lambda,\sigma,\theta,\varphi}(x - u, y - v) \, du \, dv. \end{aligned} \quad (2)$$

The model used in [6] also involves thresholding 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 (2) is known as the Gabor filter.

2.2 Complex Cells and Gabor Energy Filters

The Gabor energy is related to a model of complex cells which combines the responses of a pair of simple cells with a phase difference of $\frac{\pi}{2}$. The results of a pair of symmetric and antisymmetric filters are combined, yielding the Gabor energy $E_{\lambda,\sigma,\theta}(x, y)$ as follows:

$$E_{\lambda,\sigma,\theta}(x, y) = \sqrt{r_{\lambda,\sigma,\theta,0}^2(x, y) + r_{\lambda,\sigma,\theta,-\frac{\pi}{2}}^2(x, y)}, \quad (3)$$

where $r_{\lambda,\sigma,\theta,0}(x, y)$ and $r_{\lambda,\sigma,\theta,-\frac{\pi}{2}}(x, y)$ are the outputs of a symmetric and an antisymmetric filter, respectively.

In the following, we will use Gabor energy images $E_{\lambda,\sigma,\theta_i}(x,y)$ for a number of N_θ different orientations, with θ_i given by

$$\theta_i = \frac{i\pi}{N_\theta}, \quad i = 0, 1, \dots, N_\theta - 1. \quad (4)$$

2.3 Non-CRF Inhibition

We now extend the Gabor energy operator presented above with an inhibition term to qualitatively reproduce the above mentioned non-CRF inhibition behaviour of most orientation selective cells. For a given point in the image, the inhibition term is computed in a ring-formed area surrounding the CRF centered at the concerned point, see Fig. 1(a). We use a normalized weighting function $w_\sigma(x,y)$ defined as follows:

$$w_\sigma(x,y) = \frac{1}{\|H(G_{4\sigma} - G_\sigma)\|} H(G_{4\sigma}(x,y) - G_\sigma(x,y)) \quad (5)$$

$$H(z) = \begin{cases} 0 & z < 0 \\ z & z \geq 0 \end{cases}$$

where $G_\sigma(x,y)$ is a Gaussian defined by

$$G_\sigma(x,y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2+y^2}{2\sigma^2}}. \quad (6)$$

and $\|H(G_{4\sigma} - G_\sigma)\|$ is the L_1 norm of $G_{4\sigma}(x,y) - G_\sigma(x,y)$:

$$\|H(G_{4\sigma} - G_\sigma)\| = \iint |H(G_{4\sigma}(x,y) - G_\sigma(x,y))| dx dy.$$

We model non-CRF inhibition by computing an inhibition term $t_{\lambda,\sigma,\theta_i}(x,y)$ for each orientation θ_i . This inhibition term is computed as a convolution of the Gabor energy $E_{\lambda,\sigma,\theta_i}(x,y)$ with the weighting function $w_\sigma(x,y)$:

$$t_{\lambda,\sigma,\theta_i}(x,y) = E_{\lambda,\sigma,\theta_i}(x,y) * w_\sigma(x,y) \quad (7)$$

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

$$\tilde{b}_{\lambda,\sigma,\theta_i}(x,y) = H(E_{\lambda,\sigma,\theta_i}(x,y) - \alpha t_{\lambda,\sigma,\theta_i}(x,y)), \quad (8)$$

with $H(z)$ defined as in (5). The factor α controls the strength of the inhibition of the surround on the Gabor energy operator. If there is no texture in the surrounding of a given point, the response of this operator at that point will be equal to the response of the Gabor energy operator. An edge passing through that point will be detected by the introduced operator in the same way as it is detected by the Gabor energy operator (of appropriate orientation). However, if there are other edges in the surrounding, the inhibition term $t_{\lambda,\sigma,\theta_i}(x,y)$ may become so strong that it cancels completely the contribution of the Gabor energy operator, resulting in zero response of the operator

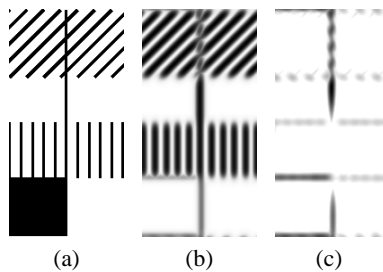


Fig. 2. (a) Synthetic input image. (b) The Gabor energy operator responds to lines and edges independently of the context, i.e., the surrounding in which these lines and edges are embedded. (c) The bar cell operator responds selectively to isolated lines and edges and lines that are surrounded by a grating of a different orientation. It does not respond to bars which make part of a grating.

introduced above. Notice that edges of the same orientation as the main stimulus will have stronger suppression effect than edges of different orientations. We refer to this type of suppression as *anisotropic* inhibition.

We now construct a new map $b_{\lambda,\sigma}(x, y)$ with values of maximum response of $\tilde{b}_{\lambda,\sigma,\theta_i}(x, y)$ over all orientations:

$$b_{\lambda,\sigma}(x, y) = \max\{\tilde{b}_{\lambda,\sigma,\theta_i}(x, y) \mid i = 0, 1, \dots, N_\theta - 1\}, \quad (9)$$

and an orientation map $\Theta(x, y)$ with the orientation for which this maximum response is achieved:

$$\begin{aligned} \Theta(x, y) &= \theta_k, \\ k &= \operatorname{argmax}\{\tilde{b}_{\lambda,\sigma,\theta_i}(x, y) \mid i = 0, 1, \dots, N_\theta - 1\}. \end{aligned} \quad (10)$$

The orientation map is needed in the binarization post-processing step, which involves nonmaxima suppression and hysteresis thresholding [10]. The operator defined by (9) will respond to isolated lines, edges, and bars, but it will not respond to groups of such stimuli that make part of texture grating having the same orientation as a bar of optimal orientation in the CRF, see Fig. 2(c). We will refer to this operator briefly as the ‘bar cell operator’, in analogy with the function of the type of visual neuron that exhibits a similar behaviour [6].

3 Performance Evaluation

Since the bar cell operator behaves as any traditional edge detector on isolated contours, bars, and edges, but behaves differently on contours in the presence of surrounding texture, we evaluate its performance on images representing objects on textured background. We selected a set of 20 images, and for each image, an associated ground truth binary edge map was specified by a subject.

3.1 Performance Measure

Let E_{GT} and B_{GT} be the set of edge pixels and background pixels of the ground truth edge image, respectively, and E_D and B_D be the set of edge pixels and background pixels of the operator-detected edge image, respectively. The set of correctly detected edge pixels is $E = E_D \cap E_{GT}$. False negatives, i.e. ground-truth edges missed by the edge detector, are given by the set $E_{FN} = E_{GT} \cap B_D$, while false positives (spurious edges) are given by the set $E_{FP} = E_D \cap B_{GT}$. We define the following measure of the *performance* of the edge detector as:

$$P = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FP}) + \text{card}(E_{FN})}, \quad (11)$$

in which $\text{card}(X)$ denotes the number of elements of set X . The performance measure P is a scalar taking values in the interval $[0, 1]$. If all true edge pixels are correctly detected and no background pixels are falsely detected as edge pixels, then $P = 1$. For all other cases, the performance measure takes values smaller than one, being closer to zero as more edge pixels are falsely detected and/or missed by the edge detector operator. Since edges cannot always be detected at exact integer image coordinates, we consider that an edge pixel is correctly detected if a corresponding ground truth edge pixel is present in a 5×5 square neighborhood centered at the respective pixel coordinates.

3.2 Experimental Results

We assess the performance of the bar cell contour detector, and compare it to the performance of the Canny edge detector [10]. This operator involves gradient magnitude computation followed by nonmaxima suppression and hysteresis thresholding [10]. We compute the gradient magnitude by convolution with the two directional derivatives of a Gaussian, as proposed in [11].

The Canny edge detection operator has three parameters: (1) σ , the standard deviation of a Gaussian smoothing kernel, (2) p , a percentage of candidate edge pixels which are retained in the final edge map, from which a corresponding threshold value t_h is computed, and (3) t , a fraction of t_h . The parameters p and t are used by hysteresis thresholding. The bar cell contour detector has an additional parameter α , which is the texture suppression factor.

For the Canny edge detector, we used eight scales, $\sigma \in \{1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4\}$. For the bar cell contour detector, we used four scales covering the same domain, $\sigma \in \{1.2, 1.6, 2.0, 2.4\}$ and two texture suppression factors, $\alpha \in \{1.0, 1.2\}$. For both methods, we applied five values of p , $p = \{50\%, 40\%, 30\%, 20\%, 10\%\}$; t was fixed to $t = 0.5 t_h$. This results in 40 parameter combinations for each of the methods. We used $N_\theta = 12$.

Figure 3 shows edge maps for some of our test images. The first and second column show the input image and the ground truth image, respectively, while the third, and fourth column show the results for the Canny edge detector and the bar cell contour detector, respectively. The result images shown in Figure 3 correspond to parameter values for which best results were obtained according to the performance measure

introduced in (11). We also computed the percentage of false positives e_{fp} as the number of false positives divided by the number of correctly detected edge pixels ($e_{fp} = \text{card}(E_{FP})/\text{card}(E)$), and the percentage of false negatives e_{fn} as the number of false negatives divided by the number of ground truth edge pixels ($e_{fn} = \text{card}(E_{FN})/\text{card}(GT)$). The performance measures and percentages of false positives and false negatives are displayed below each image.

The results show that the bar cell contour detector indeed suppresses edges in the presence of surrounding texture. The second row (Goat 3) and third row (Gazelle) are very good examples of this behaviour. The example on the second row (Goat 3) shows also that the Canny edge map contains so many spurious edges that it is hard to distinguish the contours of the object. In contrast, distinguishing the object contours in the bar cell edge map can be done without difficulty. The performance measure is consistently higher for the bar cell contour detector (Fig. 4), and this is mostly due to a reduced percentage of false positives. This is in agreement with the proposed model of the bar cell: edges resulting from a texture background are suppressed, and object contours are retained. An interesting case is the synthetic image on the fourth row. The response of the bar cell operator agrees with our perceptual interpretation of the image: two lines superimposed on oriented texture.

4 Discussion

The non-CRF inhibition algorithm presented in this paper treats different classes of edges and lines in different ways: *single* contour lines and edges are not effected by the inhibition, since they are considered as non-texture features, while *groups* of lines and edges are suppressed, since they are viewed as texture features. The model of non-CRF inhibition we use in this study is simple and straightforward: the response of an orientation and scale specific operator in a given position is suppressed by the responses of the same operator in other positions outside the CRF. The model makes use of a single parameter α that is the coefficient with which the weighted summation inhibition term is taken into account. The value of this parameter can be determined in an optimization problem derived from a specific goal, e.g. maximization of the performance of the operator for a certain set of images. Actually, for almost all natural images used we obtained the best results with a value of α around 1. One can think of more intricate or general non-CRF inhibition models, such as the division normalization proposed in [12], where the inhibition weights of the responses of all orientation and scale channels and all positions are not fixed in advance but are determined such that mutual dependencies between the ultimate responses are minimized.

Inhibition mechanisms have been applied previously to biologically motivated edge detectors in order to improve certain aspects of their function. A symmetric Gabor filter, will, for instance, respond not only along a line but also alongside the line at a certain distance from it. Similarly, the largest response of an antisymmetric Gabor filter to a line will be displaced from the line. In [13–15], various inhibition mechanisms have been proposed to remove these flanking responses. These works differ from the current work in two major aspects. First, the inhibition mechanisms act within the CRF. Second,

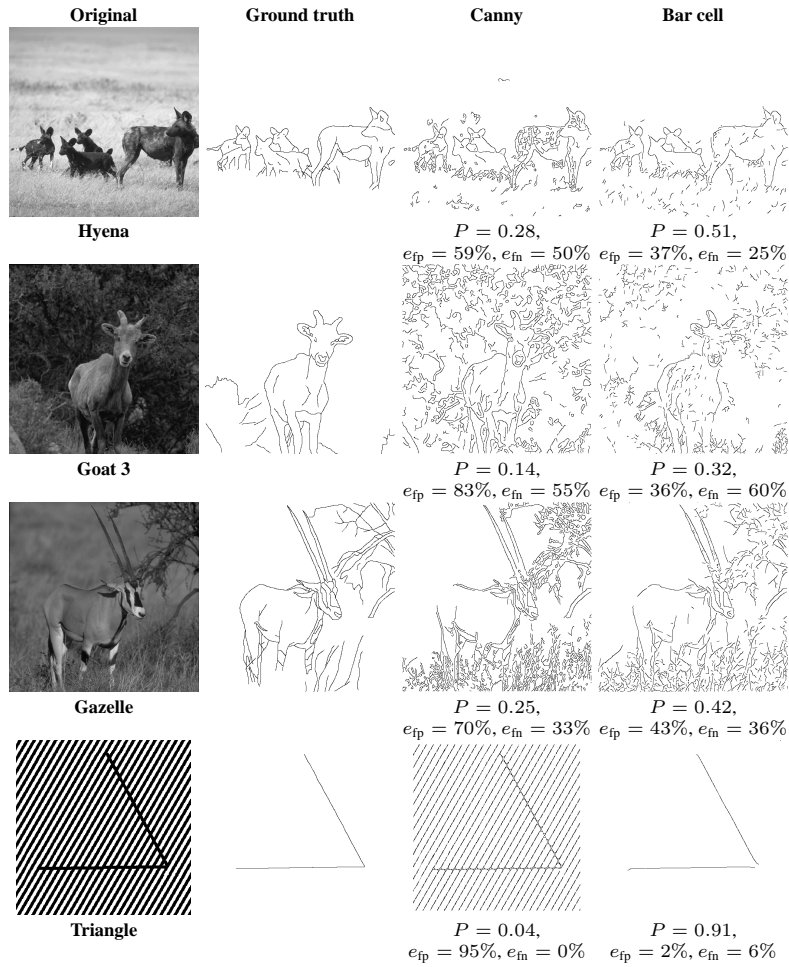


Fig. 3. Natural scenes with objects on textured background (first column), their corresponding ground truth edge maps (second column), edge maps obtained with the Canny edge detector (third column), and edge maps obtained with the bar cell contour detector (last column). Note that for a given image the best performance results (top end of the whisker) need not be obtained for the same combination of parameters of the two methods. The best performance result for the “Gazelle” image, for instance, is obtained at a finer scale (i.e. smaller value of σ) of Canny’s algorithm than the scale of the bar cell operator.

the purpose of the inhibition is quite different: it deals with the removal of flanking responses, rather than with the suppression of texture edges.

In this paper we address an anisotropic inhibition mechanism only. However, there is physiological evidence that isotropic inhibition is equally important: approximately one third of the cells with non-CRF modulation exhibit anisotropic inhibition while another third are accounted for isotropic inhibition [16]. Elsewhere [17] we studied the

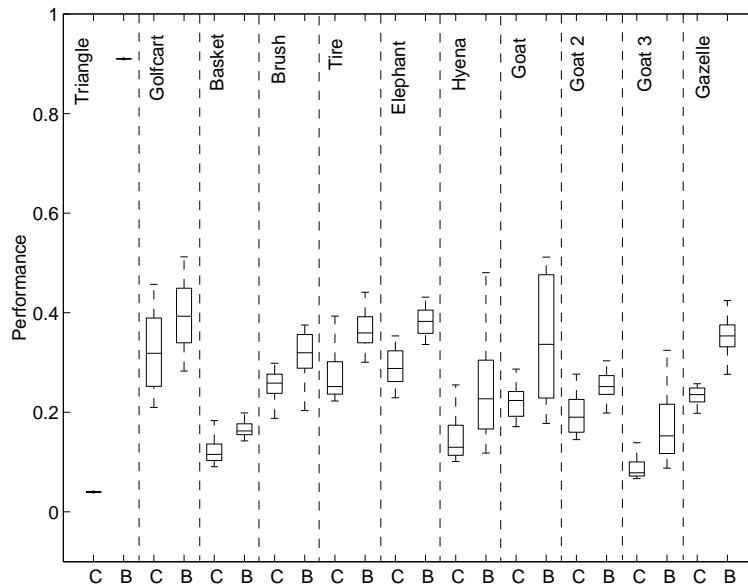


Fig. 4. Box-and-whisker plots of the performance of the Canny edge detector (denoted by C) and the bar cell contour detector with anisotropic inhibition (denoted by B) for some of the images used in our experiments.

effect of both mechanisms. For natural images like those shown in Figure 3 we obtain slightly better results using the isotropic inhibition, but perceptual effects like the ones illustrated by Figures 1(b) and (c) can only be adequately explained by anisotropic inhibition.

Applying inhibition on the Gabor energy may suggest that suppression takes place at the complex-cell level. The same inhibition mechanism can be applied to the Gabor filter output. This would correspond to suppression at the simple cell level.

The inhibitory term is computed by integrating in the surround of the response for a given preferred spatial frequency and orientation. This, however, does not mean that there will be inhibition only by gratings of that same frequency and orientation: within a certain frequency range gratings and textures of nearby orientations and frequencies will have an inhibitory effect as well. This range is determined by the band-pass characteristic of the Gabor filter [6].

Due to the limited space, various aspects could not be treated in this study. In this work we model only inhibitory modulations exhibited within the non-CRF. The physiological and psychophysical literature contains evidences for contextual facilitation as well [18]. Recent findings also show that the inhibition surround is not spatially symmetric, as suggested by our isotropic weighting function, but rather originates from particular sectors of the surround [19]. Other limitations of the current study are the

fixed size of the inhibition surround and the limited range of the inhibition factor α . All these aspects are worth undertaking further investigations.

References

1. Hubel, D.H., Wiesel, T.N.: Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *J. Physiology (London)* **160** (1962) 106–154
2. Daugman, J.G.: Uncertainty relations for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Optical Society of America A* **2** (1985) 1160–1169
3. Knierim, J.J., van Essen, D.C.: Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J. Neurophysiology* **67** (1992) 961–980
4. Schiller, P.H., Finlay, B.L., Volman, S.F.: Quantitative studies of single-cell properties in monkey striate cortex. III. spatial frequencies. *J. Neurophysiology* **39** (1976) 1334–1351
5. von der Heydt, R., Peterhans, E., Dürsteler, M.R.: Periodic-pattern-selective cells in monkey visual cortex. *J. Neuroscience* **12** (1992) 1416–1434
6. Petkov, N., Kruizinga, P.: Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells. *Biological Cybernetics* **76** (1997) 83–96
7. Blakemore, C., Carpenter, R.H.S., Georgeson, M.A.: Lateral inhibition between orientation detectors in the human visual system. *Nature* **228** (1970) 37–39
8. Kanizsa, G.: *Organization in Vision, Essays on Gestalt Perception*. Praeger, New York (1979)
9. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Analysis and Machine Intelligence* **24** (2002) 509–522
10. Canny, J.F.: A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence* **8** (1986) 679–698
11. Tagare, H., deFigueiredo, R.: On the localization performance measure and optimal edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence* **12** (1990) 1186–1190
12. Simoncelli, E.P., Schwartz, O.: Modeling surround suppression in V1 neurons with a statistically-derived normalization model. In Kearns, M.S., Solla, S.A., Cohn, D.A., eds.: *Advances in Neural Information Processing Systems 11*, Cambridge, MA, MIT Press (1999) 153–159
13. Heitger, F.: Feature detection using suppression and enhancement. Technical Report TR-163, Communication Technology Laboratory, Swiss Federal Institute of Technology (1995)
14. Petkov, N., Kruizinga, P., Lourens, T.: Lateral inhibition in cortical filters. In: *Proc. Int. Conf. on Digital Signal Processing*, Nicosia, Cyprus (1993) 122–129
15. Petkov, N., Lourens, T.: Interacting cortical filters for object recognition. In: *Proc. of the Asian Conference on Computer Vision*, Osaka, Japan (1993) 583–586
16. Nothdurft, H.C., Gallant, J., van Essen, D.: Response modulation by texture surround in primate area V1: Correlates of "popout" under anesthesia. *Visual Neuroscience* **16** (1999) 15–34
17. Grigorescu, C., Petkov, N., Westenberg, M.A.: Contour detection based on non-classical receptive field inhibition. submitted to *IEEE Trans. Image Processing* (2002)
18. Kapadia, M., Westheimer, G., Gilbert, C.: Spatial distribution of contextual interactions in primary cortex and in visual perception. *Journal of Neurophysiology* **84** (2000) 2048–2062
19. Walker, G., Ohzawa, I., Freeman, R.: Asymmetric suppression outside the classical receptive field of the visual cortex. *Journal of Neuroscience* **23** (1999) 10536–10553