NeuroImage 55 (2011) 49–66

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Center–surround interaction with adaptive inhibition: A computational model for contour detection

Chi Zeng [a], Yongjie Li [a,*], Chaoyi Li [a,b]

[a] Key Laboratory for Neuroinformation of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China
[b] Center for Life Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

ARTICLE INFO

Article history:
Received 20 August 2010
Accepted 22 November 2010
Available online 29 November 2010

Keywords:
Contour extraction
Classical receptive field
Non-classical receptive field
Surround inhibition
Adaptive end inhibition

ABSTRACT

The broad region outside the classical receptive field (CRF) of a neuron in the primary visual cortex (V1), namely non-CRF (nCRF), exerts robust modulatory effects on the responses to visual stimuli presented within the CRF. This modulating effect is mostly suppressive, which plays important roles in visual information processing. One possible role is to extract object contours from disorderly background textures. In this study, a two-scale based contour extraction model, inspired by the inhibitory interactions between CRF and nCRF of V1 neurons, is presented. The kernel idea is that the side and end subregions of nCRF work in different manners, i.e., while the strength of side inhibition is consistently calculated just based on the local features in the side regions at a fine spatial scale, the strength of end inhibition adaptively varies in accordance with the local features in both end and side regions at both fine and coarse scales. Computationally, the end regions exert weaker inhibition on CRF at the locations where a meaningful contour more likely exists in the local texture and stronger inhibition at the locations where the texture elements are mainly stochastic. Our results demonstrate that by introducing such an adaptive mechanism into the model, the non-meaningful texture elements are removed dramatically, and at the same time, the object contours are extracted effectively. Besides the superior performance in contour detection over other inhibition-based models, our model provides a better understanding of the roles of nCRF and has potential applications in computer vision and pattern recognition.

Introduction

As one of the fundamental and important tasks in object recognition related applications, contour extraction aims to isolate objects in complex scenes (Forsyth and Ponce, 2003). Computationally, extracting contour is more complicated than detecting edge. The former requires eliminating textures and distinguishing contours from non-contour edges. Such task is difficult to be well performed with traditional edge detectors such as that of Canny (1986), because they just detect all points in images where luminance changes sharply while making no distinction between contours of objects and edges originating from textured regions like grass or foliage (Forsyth and Ponce, 2003; Grigorescu et al., 2003; Papari et al., 2007). Fortunately, it has been proposed that human visual system evolved so as to be able to extract this contour feature rapidly and effectively (Albright and Stoner, 2002; Bar, 2004; Landy and Graham, 2004; Marr and Poggio, 1979), which gives us great inspiration to design physiologically plausible contour detection models.

Extensive physiological findings have clearly shown that the area beyond classical receptive field (CRF) of the majority of neurons in primary visual cortex (V1), namely non-classical receptive field (nCRF), although alone unresponsive to visual stimuli, can exert modulatory effects on the response to the stimuli in the CRF (e.g., DeAngelis et al., 1994; Hirsch and Gilbert, 1991; Hubel and Wiesel, 1968; Jones et al., 2001; Kapadia et al., 2000; Knierim and van Essen, 1992; Li, 1996; Li and Li, 1994, 1995; Walker et al., 1999, 2000; Yao and Li, 1998; for review of earlier works, see Allman et al., 1985; for a recent review, see Fitzpatrick, 2000; Series et al., 2003). These modulatory influences allow neurons to integrate information from relatively large parts of the visual field and are supposed to participate in a number of complex perceptual tasks such as contour integration and surface segmentation at the early stage of visual system (Jones et al., 2001; Kapadia et al., 2000; Li, 1996; Li and Li, 1994; Walker et al., 2000). It has been found that the modulation can be classified into two basic types, facilitation and inhibition (suppression), although the majority of the surround modulation of V1 neurons are suppressive (Guo et al., 2005; Jones et al., 2001; Li and Li, 1994; Walker et al., 2000). Our previous electrophysiological studies (Li, 1996; Li and Li, 1994, 1995; Yao and Li, 1998) and some other reports (DeAngelis et al., 1994; Kapadia et al., 2000; Walker et al., 1999) have shown that the modulatory influences from the side and end subregions of nCRF

* Corresponding author.
E-mail address: liyj@uestc.edu.cn (Y. Li).

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.11.067

can be similar or different for different cells. With the use of the sign "+" for facilitation, and "−" for inhibition, four possible combinations of side and end integration could be expressed as (Li, 1996; Li and Li, 1994, 1995) (A) *end*−/*side*−, (B) *end*+/*side*+, (C) *end*+/*side*−, and (D) *end*−/*side*+. The neurons in type A, for example, show inhibitory effects at both end and side subregions and vice versa for those in type B; similarly, facilitatory end regions and inhibitory side regions were found for the cells in type *C* and vice versa in type D. We have suggested that different structures of surround regions may be adapted to the analysis of different stimulus patterns (Li, 1996; Li and Li, 1994). Among the four types of combinations, type A accounts for about half (48.4%) of the total (Li, 1996; Li and Li, 1994, 1995; Yao and Li, 1998), and cells with such type of nCRF organization show weak or no response to broad field homogenous textures but respond vigorously to various textural contrasts, such as contrast in orientation (Das and Gilbert, 1999; Dragoi and Sur, 2000; Sillito et al., 1995; Somers et al., 1995), in luminance and spatial frequency (Bredfeldt and Ringach, 2002; De Valois and Tootell, 1983; Shen et al., 2007), in spatial phase (Xu et al., 2005), and in relative moving speed (Li et al., 1999). Consequently, this feature selective surround inhibition was viewed as the neural basis for perceptual 'pop out' (Dobbins et al., 1987; Kapadia et al., 1995; Knierim and van Essen, 1992; Lamme, 1995; Peterhans and von der Heydt, 1989) and for contour discrimination (De Valois and Tootell, 1983; Li and Li, 1995; Walker et al., 1999; Xu et al., 2005; Yao and Li, 1998).

Based on the insight understanding of nCRF inhibition mechanism, several nCRF inhibition-based contour detection models have recently been proposed in the literature (e.g., Grigorescu et al., 2003; La Cara and Ursino, 2008; Li, 1998; Long and Li, 2008; Papari et al., 2007; Petkov and Westenberg, 2003; Ren, 2008; Tang et al., 2007; Ursino and La Cara, 2004). Typically, Grigorescu et al. (2003) and Petkov and Westenberg (2003) proposed two models, i.e., isotropic and anisotropic inhibition models, both employ a Gabor function to represent CRF and an annular surround area to simulate nCRF. Their models enhance contour detection in cluttered visual scenes more effectively than the classical Canny edge detector. However, a close investigation shows that the underlying hypothesis of their models that stimuli in nCRF exert no suppression of edges belonging to object contours is not always true. As schematically illustrated in Fig. 1, the surround inhibition calculated by convolving the local information in the annular surround with a convolution kernel generally exhibits strong self-inhibition to the edge segment in the CRF. Although self-inhibition helps suppress texture, it is not desirable in the local image area where contour exists, because edges belonging to object contours generally traverse simultaneously the CRF and the end regions of nCRF, and hence, weak contours could be difficult to pop out because of self-inhibiting, especially for the object contours embedded in a cluttered background (Papari et al., 2007; Tang et al., 2007). In order to extract a good contour result, the models proposed by Grigorescu et al. have to keep a balance between contour self-inhibiting and texture suppressing. To avoid self-inhibition, Papari et al. (2007) split the inhibition surround into two truncated half-rings oriented along the concerned edge, and the band region of a specified width between the two half-rings contributed no surround inhibition. Such surround structure could of course help avoid self-inhibition for those isolated edges, but unfortunately, it could also partially reduce the suppression for texture regions to a certain extent. Furthermore, there is no sufficient evidence to support that spatial arrangement of nCRF is a half-ring pattern in primary visual cortex.

In this paper, we proposed a biologically motivated contour extraction model by explicitly simulating the structure of type A (*end*−/*side*−) nCRF. In this model, a butterfly-shaped surround (Fig. 2) was designed, which consists of two adaptive inhibitory end regions and two non-adaptive inhibitory side regions, with a Gabor-simulated CRF in the center. Specifically, the proposed model employs the idea of multi-resolution analysis (Mallat, 1989), and the adaptive end inhibition is realized based on the information extracted at two different spatial scales, i.e., Gabor energy information at a coarse scale and the side inhibition information at a fine scale. The general hypothesis behind our model is that coarser scales contain mainly the general morphology or object contours of images where most of the texture details (high-frequency components) disappear. Such hypothesis is thought to be reasonable based on the underlying physiological substrate of multi-resolution wavelet analysis (Campbell and Robson, 1968; Mermillod et al., 2005; Morrone and Burr, 1997; Oliva and Schyns, 1997; Papari et al., 2007; Mallat, 1989). In fact, multi-resolution mechanism has been successfully used in contour detection (Joshi and Sivaswamy, 2005; La Cara and Ursino, 2008; Liang et al., 1999; Lindeberg, 1998; Mallat, 1989; Papari et al., 2007; Ren, 2008), and this issue is further discussed in the Summary and discussion section.

The rest of this paper is organized as follows. The proposed biologically motivated model is explained in the Materials and methods section. Then experimental results and performance evaluation are presented in Results section. Finally, we discuss in the Summary and discussion section the underlying physiological relevance of our model, as well as some computational explanations why our model can produce superior results for contour detection.

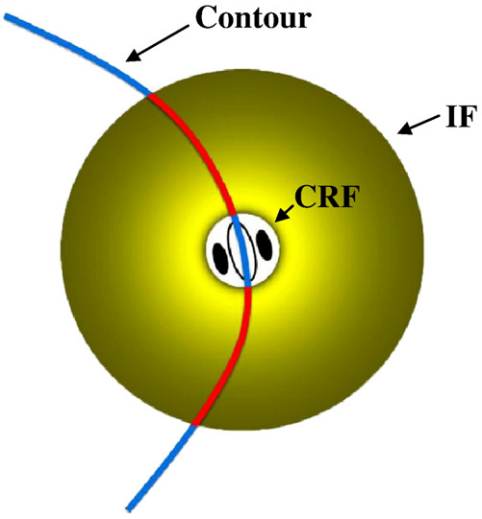


Fig. 1. Schematic drawing illustrating the self-inhibition in contour detection. The parts of contour presented in nCRF (red segments) exert inhibition to the neuronal response to the part of contour located in CRF.

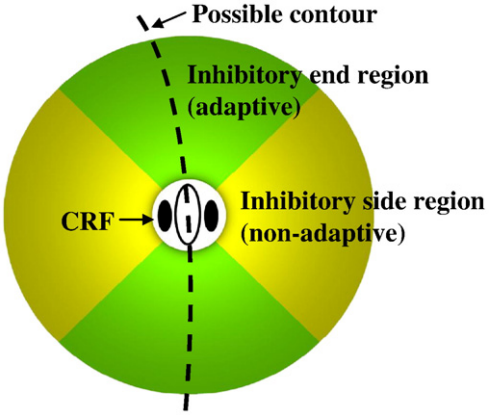


Fig. 2. The nCRF structure of the proposed model.

Materials and methods

Fig. 2 shows the schematic drawing to illustrate our nCRF model by simulating neurons in type A (*end*−/*side*−), in which the whole nCRF region is divided into four annulus sector shaped subregions (Li, 1996; Li and Li, 1994). The two annulus sectors lying along the collinear axis of CRF's preferred orientation are defined as the end regions in this study, whereas another two sectors lying along the axis perpendicular to the preferred orientation of CRF are the side regions.

As mentioned earlier, it is not a reasonable solution for contour detection to keep a consistent inhibition for nCRF without distinguishing the contextual configurations. Intuitively, a more rational alternative is to reduce the inhibition strength at the locations where object contours exist. Consequently, two questions arise: (1) how to judge whether a contour exists or not at a given location and (2) how to reasonably reduce the inhibition strength when an object contour likely exists. To demonstrate that our solutions to these two questions involved are more trackable, we start by describing the general model of center–surround interaction based contour detection in the following sections, followed by successively explaining the involved particular mechanisms, that is, the response of classical receptive field, the non-adaptive side inhibition and the adaptive end inhibition.

General model description

We simulate explicitly the visual mechanisms of contour detection based on the center–surround interaction of V1 neurons using the general model written as

$$r(x,y) = H\Big(\tilde{E}\big(x,y;\sigma_f,\lambda_f\big) - \alpha_1 \mathrm{inh}_{\mathrm{side}}\big(x,y;\sigma_f,\lambda_f\big) - \alpha_2 \mathrm{inh}_{\mathrm{end}}\big(x,y;\sigma_f,\lambda_f,\sigma_c,\lambda_c\big)\Big), \tag{1}$$

where $r(x,y)$ is the final excitatory response of a V1 neuron to the stimuli presented simultaneously in its CRF and nCRF. $H(s) = \max(s,0)$, a half-wave rectification operation simulating that the cortical cell can fire only when positive excitation is received. The neuronal response to the stimuli presented in CRF is represented as $\tilde{E}(x,y;\sigma_f,\lambda_f)$, simulating the geniculate input to the cortical cell. $\mathrm{inh}_{\mathrm{side}}(x,y;\sigma_f,\lambda_f)$ and $\mathrm{inh}_{\mathrm{end}}(x,y;\sigma_f,\lambda_f,\sigma_c,\lambda_c)$ are the side and end inhibition, respectively. α_1 and α_2 are constants denoting the intrinsic connection strengths of the side and end inhibition. (σ_f,λ_f) and (σ_c,λ_c) are Gabor-related scale parameters at a fine and coarse scale, respectively. The mathematical model given by Eq. (1) indicates that the final excitatory response is determined by the CRF response and the inhibitions originating from the side and end regions of nCRF. Among these three terms, $\tilde{E}(x,y;\sigma_f,\lambda_f)$ is simulated using a Gabor energy model; side inhibition $\mathrm{inh}_{\mathrm{side}}(x,y;\sigma_f,\lambda_f)$ is modeled in a normal (non-adaptive) way as Grigorescu et al. (2003) developed, and end inhibition $\mathrm{inh}_{\mathrm{end}}(x,y;\sigma_f,\lambda_f,\sigma_c,\lambda_c)$ is modeled in a new novel adaptive manner by utilizing the information at both the fine and coarse scales, which are described in details in the next three subsections.

Fig. 3 shows the schematic drawings illustrating the general flowchart of the proposed model for contour detection, and the solutions to these two questions involved are explained in details in the following sections. In short, given an original image (Fig. 3a), two different Gabor energy maps are calculated at a fine and coarse spatial scale (Fig. 3b and c). Consistent inhibition from the side regions is calculated based on the Gabor energy map extracted at the fine scale. In contrast, adaptive inhibition from the end regions is calculated based on the coarse-scale Gabor energy map and together with the fine-scale side inhibition and the fine-scale Gabor energy map. Note that here the term *consistent inhibition* means that the side inhibition is consistently calculated just based on the local features presented in the side regions at the fine scale, a non-adaptive manner similar to others (Grigorescu et al., 2003). In contrast, the end inhibition is adaptively weighted based on the local information presented in both end and side regions at both the fine and coarse scales.

It is worth noting that the inhibition is applied to $\tilde{E}(x,y;\sigma_f,\lambda_f)$, the Gabor energy map computed at the fine scale (σ_f,λ_f) (Fig. 3d), and the following binary contour map construction (Fig. 3e) is also performed based on the inhibited fine energy map. The necessity for such consideration comes from the fact that contours extracted at larger scales are inaccurate in localization, i.e., contours incline to be shifted away from their true locations (Liang et al., 1999; Papari et al., 2007).

Response of classical receptive field

As is well known, V1 neurons receive afferent input from neurons in the lateral geniculate nucleus (LGN), and the V1 neuronal responses to stimuli within receptive fields can be well described by the so-called Gabor function (Daugman, 1985; Jones and Palmer, 1987; McLean and Palmer, 1989; Mehrotra et al., 1992), which is a Gaussian envelope multiplied by a sinusoidal modulation. Gabor energy, namely the response modulus of orthogonal pairs of Gabor filters, can represent typically fundamental characteristics of CRF in V1 (Chan and Coghill, 2001). A two-dimensional Gabor filter can be expressed as

$$g(x,y;\sigma,\lambda,\varphi,\theta) = e^{-\frac{\tilde{x}^2+r^2\tilde{y}^2}{2\sigma^2}} \cos\left(2\pi\frac{\tilde{x}^2}{\lambda}+\varphi\right), \tag{2}$$

where $\tilde{x} = x\cos\theta + y\sin\theta$, $\tilde{y} = -x\sin\theta + y\cos\theta$, in which θ is CRF's preferred orientation. φ is a phase offset, and $g(x,y;\sigma,\lambda,\varphi,\theta)$ with $\varphi=0$ and $\varphi=\pi/2$ represent the even and odd Gabor filters, respectively. r is the spatial aspect ratio that determines the eccentricity of the Gaussian envelope. We set $r=0.5$ in this work, since it has been found that the aspect ratio varies in the range 0.23 (nearly 5 to 1 elongation) to 0.92 (nearly round) for those simple receptive fields in cat striate cortex (Jones and Palmer, 1987). σ is the standard deviation of Gaussian factor that determines the size of the Gabor-simulated CRF. λ is the wavelength. σ/λ represents the spatial frequency bandwidth. Since most of the cells in visual cortex have bandwidths between 1.0 and 1.5 octaves (De Valois et al., 1982), we fix σ/λ in this study to the value 0.56, a biologically plausible value that corresponds to a bandwidth of about one octave at half-response (Kruizinga and Petkov, 1999) and that has been commonly used in Gabor-based visual system models (Grigorescu et al., 2003; Kruizinga and Petkov, 1999; Petkov and Westenberg, 2003). Note that since λ and σ are dependent with each other by setting $\sigma/\lambda=0.56$, we choose σ as a free parameter in the simulations of this study.

Neuronal response $E(x,y;\sigma,\lambda,\theta)$ to the stimuli in CRF is given by

$$E(x,y;\sigma,\lambda,\theta) = \sqrt{e(x,y;\sigma,\lambda,0,\theta)^2 + e(x,y;\sigma,\lambda,-\pi/2,\theta)^2}, \tag{3}$$

$$e(x,y;\sigma,\lambda,\varphi,\theta) = I(x,y) * g(x,y;\sigma,\lambda,\varphi,\theta), \tag{4}$$

where $I(x,y)$ represents the stimuli (i.e., input image) and $*$ denotes the convolution operator to combine the input to a cortical cell coming from all surrounding cells. $e(x,y;\sigma,\lambda,0,\theta)$ and $e(x,y;\sigma,\lambda,-\pi/2,\theta)$ are the responses of even and odd Gabor-simulated cells, respectively. Gabor energy maps are computed at a number of N_θ different orientations:

$$\theta_i = \frac{(i-1)\pi}{N_\theta}, i = 1,2,\cdots,N_\theta. \tag{5}$$

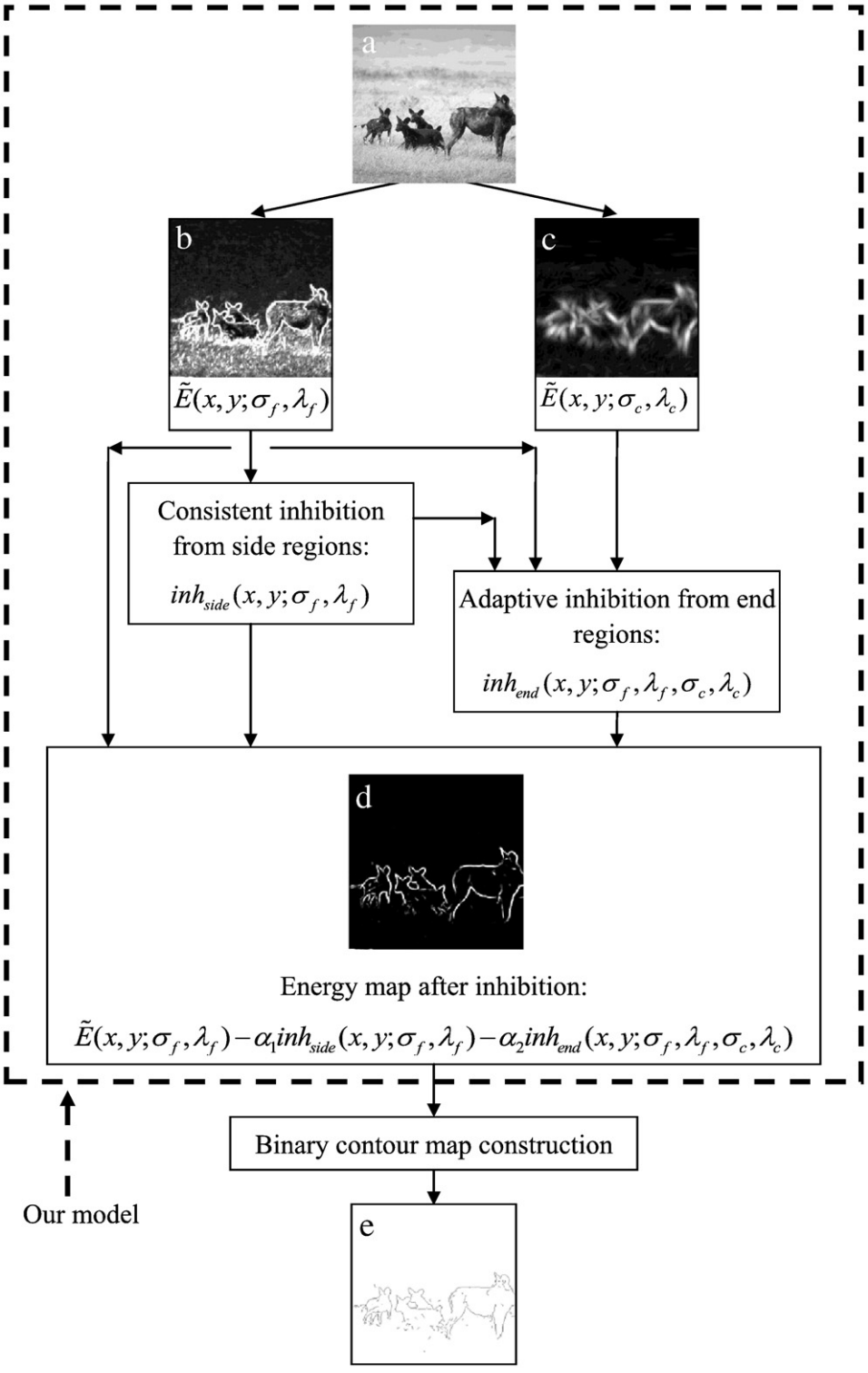


Fig. 3. Schematic drawings illustrating the general flowchart of the proposed model for contour detection as shown in the dashed box. (a) Original image. (b) The maximum Gabor energy map $\tilde{E}(x,y;\sigma_f,\lambda_f)$ at a fine scale (σ_f,λ_f), with which the inhibition from side regions $\mathrm{inh}_{\mathrm{side}}(x,y;\sigma_f,\lambda_f)$ is computed. (c) The energy map $\tilde{E}(x,y;\sigma_c,\lambda_c)$ at a coarse scale (σ_c,λ_c), with which, together with $\tilde{E}(x,y;\sigma_f,\lambda_f)$ and $\mathrm{inh}_{\mathrm{side}}(x,y;\sigma_f,\lambda_f)$, the inhibition from end regions $\mathrm{inh}_{\mathrm{end}}(x,y;\sigma_f,\lambda_f,\sigma_c,\lambda_c)$ is adaptively computed. (d) The energy map after inhibition, which is computed by subtracting the side inhibition $\mathrm{inh}_{\mathrm{side}}(x,y;\sigma_f,\lambda_f)$ and the end inhibition $\mathrm{inh}_{\mathrm{end}}(x,y;\sigma_f,\lambda_f,\sigma_c,\lambda_c)$ from the Gabor energy map $\tilde{E}(x,y;\sigma_f,\lambda_f)$ defined at the fine scale. (e) The binary contour map constructed by a process of binarization based on the inhibited energy map. In this work, we employ the binarization method of nonmaxima suppression followed by hysteresis thresholding (Canny, 1986). Note that this binarization is not an essential part of our model, and it can be replaced by any other binarization techniques. Also note that the responses in (b), (c), and (d) are normalized into the grey levels with a range of 0 (black) to 255 (white). Higher grey levels correspond to stronger responses.

The maximum Gabor energy $\tilde{E}(x,y;\sigma,\lambda)$ and preferred orientation map $\tilde{\theta}(x,y;\sigma,\lambda)$ of CRF are consequently given by

$$\tilde{E}(x,y;\sigma,\lambda) = \max\{E(x,y;\sigma,\lambda,\theta_i)\,|\,i=1,2,\cdots,N_\theta\}, \tag{6}$$

$$\tilde{\theta}(x,y;\sigma,\lambda) = \arg\max\{E(x,y;\sigma,\lambda,\theta_i)\,|\,i=1,2,\cdots,N_\theta\}. \tag{7}$$

In this work, we choose 12 different orientations with equal intervals, i.e., set $N_\theta=12$ for Eq. (5). We have found that selection of too many orientations only contribute a slight improvement in contour detection.

Surround inhibition model

We start by assuming that the overall surround inhibition to the CRF is the summation of the individual inhibitions from the four separate subregions (Li, 1996; Li and Li, 1995). In addition, as mentioned earlier, orientation-selective inhibition and other visual feature selective modulations are important underlying mechanisms for visual information processing. For simplicity, in this study, we only consider the isotropic inhibition (Grigorescu et al., 2003; Petkov and Westenberg, 2003), i.e., nCRF suppression is independent of their preferred orientations. This work focuses on exploring the adaptivity of nCRF modulation for contour detection, and the proposed scheme could be easily extended to orientation selectivity based anisotropic inhibition.

Sufficient physiological findings indicate that inhibitory strength from nCRF decreases with increasing distance from the center of the RF (Chen et al., 2005; Li, 1996; Li and Li, 1994, 1995; Jones et al., 2001; Kapadia et al., 2000; Nothdurft et al., 1999; Rossi et al., 2001; Walker et al., 2000). Similar as other investigators (Papari et al., 2007; Ren, 2008; Tang et al., 2007), in this study, we adopt the inhibitory weighting model established by Grigorescu et al. (2003), while modifying it to work only in a restricted regions, i.e., side or end regions. For a given point (x,y) in the image, the weighting function is defined as

$$W_D(x,y;\sigma,k) = \frac{\mathrm{DOG}^+(x,y;\sigma,k)}{\|\mathrm{DOG}^+(x,y;\sigma,k)\|_1}, \tag{8}$$

where $\|\cdot\|_1$ denotes the L_1 norm and $\mathrm{DOG}^+(x,y;\sigma,k)$ is calculated as

$$\mathrm{DOG}^+(x,y;\sigma,k) = H\left(\frac{1}{2\pi(k\sigma)^2}\exp\left(-\frac{x^2+y^2}{2(k\sigma)^2}\right) - \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)\right). \tag{9}$$

where $H(s)=\max(s,0)$. Thus, $\mathrm{DOG}^+(x,y;\sigma,k)$ defines the ring-shaped surround region for nCRF by only taking the positive part of the difference-of-Gaussian (DOG) function. Computationally, the weighting function $W_D(x,y;\sigma,k)$ simulates the strength of the intracortical synapses linking the surrounding cells to a cortical cell (Ursino and La Cara, 2004). k is the ratio of the standard divisions of two Gaussian functions. In other word, it represents the size ratio of nCRF to CRF. In this paper, we set $k=4$ based on the neurophysiological findings that the size of nCRF is normally 2 to 5 times larger (in diameter) than that of CRF (Kapadia et al., 2000; Maffei and Fiorentini, 1976; Li and Li, 1994).

Consistent inhibition from side regions

Considering that those texture details occur most likely at relatively fine scales, we compute the side inhibition at a fine scale (σ_f,λ_f) in order to capture the texture details sufficiently and suppress them significantly. Thus, the inhibition coming from the side regions can be computed as a convolution of the Gabor energy $\tilde{E}(x,y;\sigma_f,\lambda_f)$ with a weighting function $W_{\mathrm{side}}(x,y)$:

$$\mathrm{inh}_{\mathrm{side}}\left(x,y;\sigma_f,\lambda_f\right) = \tilde{E}\left(x,y;\sigma_f,\lambda_f\right) * W_{\mathrm{side}}(x,y). \tag{10}$$

For the side regions, the weighting function $W_{\mathrm{side}}(x,y)$ is given by

$$W_{\mathrm{side}}(x,y) = w_s(x,y)W_{\mathrm{DOG}}\left(x,y;\sigma_f,k\right). \tag{11}$$

$$w_s(x,y) = \begin{cases} 1, & |\beta|\le\Phi \\ 0, & \text{else} \end{cases}, \tag{12}$$

where $w_s(x,y)$ defines the spatial range of the side regions, β represents the angle between the preferred orientation of CRF and the line from the point (x,y) to the center, Φ is the central angle of a sector shaped side region. In this model, we divide the whole nCRF region into four annulus sector shaped side and end regions with equal central angles, which results in $\Phi=\pi/4$, because there is no sufficient evidence found to date to demonstrate that any subregion is obviously larger than others.

It is worth noting that the final contour results are not very sensitive to the parameters (σ_f,λ_f), because most of fine textures in a image could be captured by the Gabor filters with a large range of fine scales. This non-sensitivity will be demonstrated by the simulation results shown in the Results section.

Adaptive inhibition from end regions

Similar to the side inhibition defined by Eq. (10), the conventional non-adaptive end inhibition could be given by

$$\mathrm{inh}'_{\mathrm{end}}\left(x,y;\sigma_f,\lambda_f\right) = \tilde{E}\left(x,y;\sigma_f,\lambda_f\right) * w_{\mathrm{end}}(x,y), \tag{13}$$

$$W_{\mathrm{end}}(x,y) = (1-w_s(x,y))W_D\left(x,y;\sigma_f,k\right), \tag{14}$$

where $W_{\mathrm{end}}(x,y)$ defines the distance related weighting template for the two end subregions. $(1-w_s(x,y))$ determines the spatial range of end subregions. $w_s(x,y)$ and $W_D(x,y;\sigma_f,k)$ are given by Eqs. (12) and (8), respectively.

As mentioned earlier, the consistent inhibition defined as Eq. (13) without distinguishing contextual information is not reasonable for contour detection. The underlying objective of adaptive end inhibition in our computational model is to exert stronger inhibition to textured areas and weaker inhibition to those potential object contours and thus lead to a significant elimination of textures while retaining contours by the greatest extent. In our computational model, we formulate the adaptive end inhibition as

$$\mathrm{inh}_{\mathrm{end}}\left(x,y;\sigma_f,\lambda_f,\sigma_c,\lambda_c\right) = \mathrm{inh}'_{\mathrm{end}}\left(x,y;\sigma_f,\lambda_f\right)W_{\mathrm{adp}}\left(x,y;\sigma_f,\lambda_f,\sigma_c,\lambda_c\right), \tag{15}$$

where $W_{\mathrm{adp}}(x,y;\sigma_f,\lambda_f,\sigma_c,\lambda_c)$ is a location-dependant weighting factor that varies adaptively in accordance with the local contextual features occurring at both of the fine scale (σ_f,λ_f) and coarse scale (σ_c,λ_c). The definition of $W_{\mathrm{adp}}(x,y;\sigma_f,\lambda_f,\sigma_c,\lambda_c)$ is explained as follows, and we simplify it as $W_{\mathrm{adp}}(x,y)$.

Intuitively, it is rational to suppose that if there are many stimuli presented in the side regions of a cell's nCRF, this cell is most likely stimulated by a textural patch. Furthermore, as mentioned earlier, most of the Gabor energy at coarse scale is actually contributed by the luminance changes at object boundaries. So, for a cell whose RF is centered at (x,y), we assume that texture regions generally contribute stronger side inhibition $\mathrm{inh}_{\mathrm{side}}(x,y;\sigma_f,\lambda_f)$ at the fine scale (σ_f,λ_f) than that at the coarse scale (σ_c,λ_c), and meanwhile, object contours produce relatively larger Gabor energy $\tilde{E}(x,y;\sigma_c,\lambda_c)$ at the coarse scale (σ_c,λ_c) than that of cluttered textures at the same coarse scale. According to the multi-resolution information processing theory and the descriptive analysis mentioned above, these assumptions are generally reasonable although they do not always hold for some quite cluttered scenes. With these assumptions, we subsequently define the following adaptive weight function for end inhibition to quantitatively

reflect the combined inhibition contributions of the object contour and the textures among the local features at coarse and fine scales:

$$W_{\mathrm{adp}}(x,y) = 1 - W_c(x,y;\sigma_c,\lambda_c) + W_f\left(x,y;\sigma_f,\lambda_f\right), \tag{16}$$

$$W_c(x,y;\sigma_c,\lambda_c) = f_{\mathrm{sig}}\left(\frac{\tilde{E}(x,y;\sigma_c,\lambda_c)}{||\tilde{E}(x,y;\sigma_c,\lambda_c)||_1}\right), \tag{17}$$

$$W_f\left(x,y;\sigma_f,\lambda_f\right) = f_{\mathrm{sig}}\left(\frac{\mathrm{inh}_{\mathrm{side}}\left(x,y;\sigma_f,\lambda_f\right)}{||\mathrm{inh}_{\mathrm{side}}\left(x,y;\sigma_f,\lambda_f\right)||_1}\right), \tag{18}$$

$$f_{\mathrm{sig}}(t) = 1/\left(1+e^{-a(t-\tau)}\right), \tag{19}$$

where $f_{\mathrm{sig}}(t)$ is a sigmoid function, which slope and horizontal shift are controlled by parameter a and τ, respectively. The right-side hand of Eq. (16) includes three items. Substituting Eq. (16) into Eq. (15), we can clearly see that these three items represent the conventional isotropic end inhibition, the reduced inhibition suggested by coarse-scale information, and the strengthened inhibition suggested by the stimulus information within the side regions. The relationships with biological evidence under these three items will be discussed in detail in Biological plausibility and a possible neural network section, where a possible neural network is proposed on the basis of underlying physiological evidences.

It should be clear from Eqs. (16)–(19) that a weaker Gabor energy $\tilde{E}(x,y;\sigma_c,\lambda_c)$ results in a smaller $W_c(x,y;\sigma_c,\lambda_c)$, and a stronger side inhibition $\mathrm{inh}_{\mathrm{side}}(x,y;\sigma_f,\lambda_f)$ results in a higher $W_f(x,y;\sigma_f,\lambda_f)$; both result in a stronger $W_{\mathrm{adp}}(x,y)$ and hence strengthen the end inhibition-based on Eq. (15). In contrast, a larger $\tilde{E}(x,y;\sigma_c,\lambda_c)$ and a smaller $\mathrm{inh}_{\mathrm{side}}(x,y;\sigma_f,\lambda_f)$ produce a smaller $W_{\mathrm{adp}}(x,y)$ and hence weaken the end inhibition. According to the assumptions mentioned earlier, if we apply a stronger end inhibition when $W_{\mathrm{adp}}(x,y)$ is larger, we can suppress the textures more significantly at a quite lower risk of destroying object contours due to the lower chance for contours to exist at such situation. Similarly, if we reduce the end inhibition when $W_{\mathrm{adp}}(x,y)$ is smaller, the chance of being retained for object contours will greatly increase.

Note that the sigmoid curve is supposed to be able to well describe the activation function for the excitatory cells and has been widely used in biologically inspired neural networks (Haykin, 1998; Li, 1998). Mathematically, this S-shaped transfer function produces no activation for inputs smaller than the threshold τ, little activation for weak inputs, nearly linear activation for intermediate inputs, and inclines to saturate quickly for those high inputs. Intuitively, such generalization helps eliminate weak noise and normalize the activities for too high inputs. In this study, we experimentally set $a=40$ and $\tau=0.25$. It is worth noting that, considering that we normalize the input to sigmoid function to the range [0, 1], these parameter settings for the sigmoid activation function are close to the ranges tested based on electrophysiological measures (Marreiros et al., 2008; Moran et al., 2007).

Binary contour map construction

To simplify comparison, just same as Grigorescu et al. (2003), we construct the binary contour maps using the standard procedure of nonmaxima suppression followed by hysteresis thresholding (Canny, 1986). Based on the response $r(x,y)$ computed using Eq. (1) and the corresponding orientation map $\tilde{\theta}(x,y)$ given by Eq. (7), the edges are thinned to one-pixel-wide candidate contours by nonmaxima suppression, and the final binary contour map is extracted from the candidates using hysteresis thresholding. Note that in the process of hysteresis thresholding, we fix the two involved threshold values t_{Bh} and t_{Bl} to $t_{\mathrm{Bl}}=0.5t_{\mathrm{Bh}}$, the same values as those of Grigorescu et al. (2003).

Results

In this section, we first give experimental results on a synthetic image to demonstrate in which situation that our model offers advantages over previous methods. Then we test our model with a population of 40 real images, and with the mean of box-and-whisker diagrams, quantitative performance is compared with existing models based on the ground truth contour maps.

Inhibitory effects

To better understand and evaluate the inhibitory behavior of the proposed model, we study a synthetic image to illustrate the results at each process step, as shown in Fig. 4. In the original image shown in Fig. 4a, the background is composed of pseudorandomly orientated short bars with two different intensity levels, in which three well-organized structures are embedded, i.e., a long vertical straight line, a short oblique straight line, and a question mark shaped curve. It is clear that the upper part of the long vertical line is the one that could be most easily detected because of the high contrast in luminance (see details in the caption of Fig. 4). Fig. 4b and c shows the Gabor energy maps computed using Eq. (6) at a fine scale $\sigma_f=1.5$ and a coarse scale $\sigma_c=6.0$, respectively. Just as expected, when the Gabor-simulated CRF of a model cell has a fine scale, both the background texture and the organized structures are almost equally responded by the model cell with only CRF (Fig. 4b). In contrast, the organized structures are much better responded by the model cell whose CRF has a coarse scale (Fig. 4c). Note that we take the fine Gabor energy map as the reference map, on which the inhibition operation is performed for all the contour detection models mentioned below.

To compare the inhibitory effects of different models, we first re-implement the anisotropic and isotropic inhibition models of Grigorescu et al. (2003). Fig. 4d and e shows, respectively, the energy map inhibited using anisotropic model with the inhibition strength factor $\alpha=3.0$ and $\alpha=4.0$ (see Eq. (8) in Grigorescu et al., 2003). Fig. 4f and g shows, respectively, the energy map inhibited using isotropic model with the inhibition strength factor $\alpha=3.0$ and $\alpha=4.0$ (see Eq. (14) in Grigorescu et al., 2003). From the figures, we can find that the anisotropic inhibition, an orientation-selective based operator, is difficult to suppress the randomly orientated bars in the lower part of the image due to the relatively lower orientation contrast in the background. Furthermore, with the increase of the inhibition strength factor from 3.0 to 4.0, the organized structures (e.g., the oblique straight line) are more heavily suppressed than the background, which could be explained by the strong self-inhibition resulting from the adopted anisotropic inhibition strategy that a less orientation difference between the CRF and the surround produces a stronger inhibition (Grigorescu et al., 2003; Petkov and Westenberg, 2003). The self-inhibition also exists for the isotropic model, as shown in Fig. 4f and g, but the organized structures and the background are equivalently suppressed due to the orientation-independent inhibition. However, the organized structures are still difficult to pop out even if the inhibition strength factor increases from 3.0 (Fig. 4f) to 4.0 (Fig. 4g).

Fig. 4h and i shows the strength of original side and end inhibition in our model computed using Eqs. (10) and (13), respectively. Taking the organized structure of long vertical line as example, we can clearly see that its side inhibition is very weak due to the sparse background bars (Fig. 4h). In contrast, the end inhibition shown in Fig. 4i is very strong, because the cell's end subregions are covered by the organized structure (see graphical illustration in Fig. 2). There is no doubt that the organized structure would be remarkably suppressed if this end inhibition was directly exerted on the response of central CRF shown

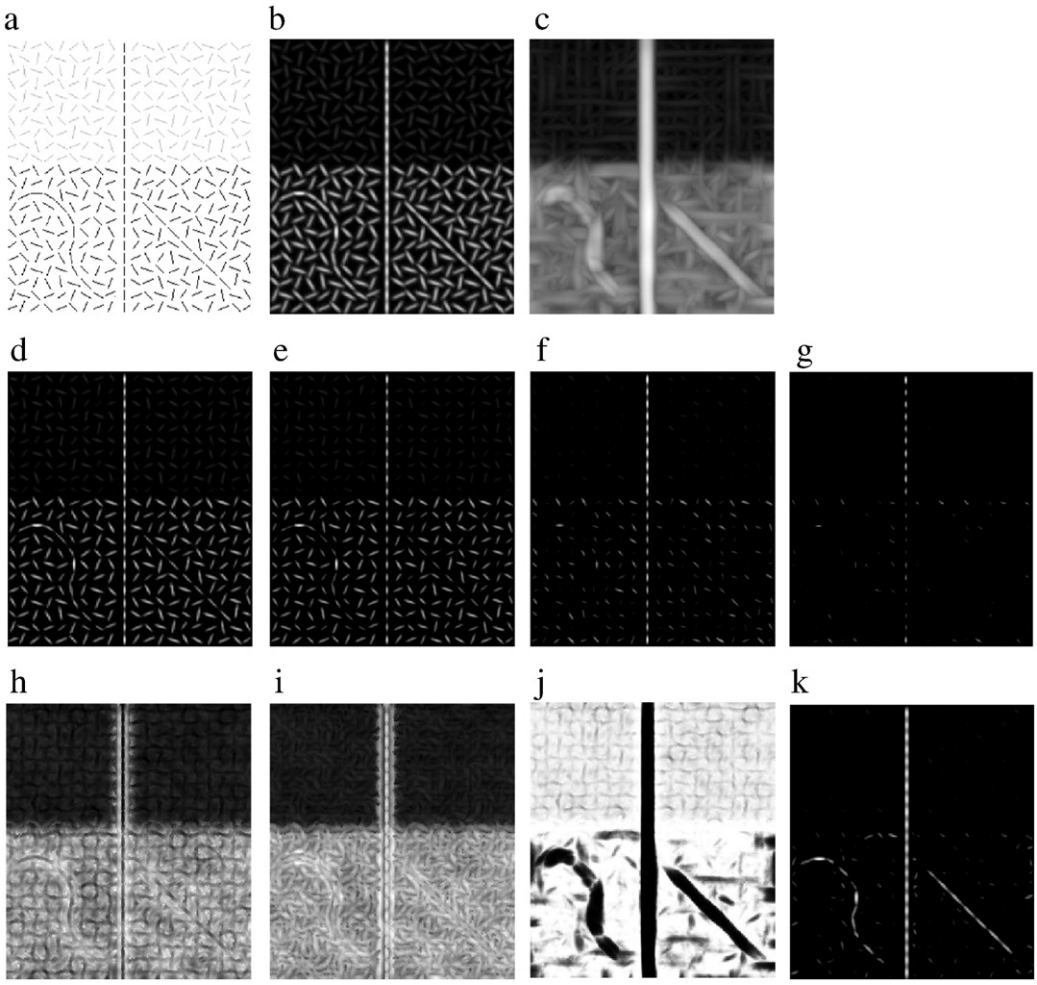


Fig. 4. Results of surround inhibition on a synthetic image. (a) Original image. (b) The maximum Gabor energy map at the fine scale with $\sigma_f = 1.5$. (c) The maximum Gabor energy map at the coarse scale with $\sigma_c = 6.0$. (d) Gabor energy map inhibited using anisotropic model with the inhibition strength factor $\alpha = 3.0$ (Eq. (8) in Grigorescu et al., 2003). (e) Gabor energy map inhibited using anisotropic model with $\alpha = 4.0$ (Grigorescu et al., 2003). (f) Gabor energy map inhibited using isotropic model with $\alpha = 3.0$ (Eq. (1414) in Grigorescu et al., 2003). (g) Gabor energy map inhibited using isotropic model with $\alpha = 4.0$ (Grigorescu et al., 2003). Note that the inhibitions in (d) to (g) are applied on the fine Gabor energy map shown in (b). (h) The strength of side inhibition in our model computed using Eq. (10). (i) The strength of end inhibition in our model computed using Eq. (13). (j) The end inhibition weighting factor computed using Eq. (17). (k) The final inhibited Gabor energy map of our model computed using Eq. (1) with $\alpha_1 = \alpha_2 = 4.0$. Note that except (a) and (j), the Gabor energy maps in all the panels are normalized into the grey levels with a range of 0 (black) to 255 (white). Higher grey levels correspond to stronger energies. The grey values of the randomly orientated short bars are 180 (lighter) and 0 (black) for the upper and lower parts of (a), respectively. The end inhibition weighting factor shown in (j) is in the range of 0.0 (black) to 1.0 (white).

in Fig. 4b. To avoid this unwanted self-inhibition, we compute the end inhibition weighting factor using Eqs. (16)–(19), which combines the information of the coarse Gabor energy (Fig. 4c) and the side inhibition (Fig. 4h). Intuitively, the local regions with stronger coarse Gabor energy reflect the higher possibility of the existence of contours. In addition, the cluttered background normally corresponds to the regions with high side inhibition. Consequently, the local regions with possible contours have lower end inhibition weighting factor, as shown by the darker regions in Fig. 4j. By combining the side inhibition (Fig. 4h) and the weighted end inhibition (Fig. 4i) as the surround modulation of CRF responses as the fine scale shown in Fig. 4b, we successful make the organized structures pop out by adaptively suppressing the background (Fig. 4k).

Contour extraction on real images

Some natural images are employed to verify the performance of our model for contour extraction, and the real images used in this study are the whole RuG dataset of 40 images downloaded from http://www.cs.rug.nl/~imaging/, a database that has been widely used to evaluate the performance of contour detectors based on the associated ground truth contour maps drawn by hand (Grigorescu et al., 2003; Papari et al., 2007; Ren, 2008; Tang et al., 2007). The performance is compared with that of two typical surround inhibition-based contour operators of Petkov and Westenberg (2003) and Grigorescu et al. (2003).

To fully evaluate the robustness of the proposed contour detector, we test it with various combinations of specific model parameters. For the Gabor filters, we use $\sigma_f \in \{1.2, 1.6, 2.0, 2.4\}$ and $\sigma_c \in \{5\sigma_f, 6\sigma_f\}$ as the fine and coarse scales. For the side and end attenuation factors defined in Eq. (1), we try them as $\alpha_1 = 1.0$ and $\alpha_2 \in \{1.0, 1.2\}$. For the binary contour map construction using nonmaxima suppression and hysteresis thresholding, we use $p \in \{0.5, 0.4, 0.3, 0.2, 0.1\}$, which defines the minimum fraction of candidate pixels to be retained in the final contour map. To compare our model with the two inhibition models of Grigorescu et al. (2003), we re-implement their models and test them with the parameters $\alpha \in \{1.0, 1.2\}$ and $\sigma \in \{1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4\}$. Totally, all of the three models are tested with 80 groups of different parameter combinations for each image.

The performance measure proposed by Grigorescu et al. (2003) is used to quantitatively evaluate the detection performance. Given the ground truth contour maps, let E_{GT} denote the set of contour pixels of a binary ground truth contour map, E denote the set of correctly

detected contour pixels, E_{FP} denote the set of false positives (spurious contours), and E_{FN} denote the set of false negatives, i.e., the pixels of ground-truth contours missed by the contour detector. The percentage of false positives e_{FP}, the percentage of false negatives e_{FN}, and the overall performance measure P are computed as

$$e_{FP} = \text{card}\,(E_{FP}) / \text{card}\,(E), \tag{20}$$

$$e_{FN} = \text{card}\,(E_{FN}) / \text{card}\,(E_{GT}), \tag{21}$$

$$P = \frac{\text{card}\,(E)}{\text{card}\,(E) + \text{card}\,(E_{FP}) + \text{card}\,(E_{FN})}, \tag{22}$$

where card (S) denotes the number of elements of the set S. It is clear from the definitions that a lower measure e_{FP} means a better suppression of texture, whereas a smaller e_{FN} indicates a better retaining of object contour, and as a whole, a higher measure P corresponds to a better overall performance.

Fig. 5 shows the best results for four images selected from the image dataset tested with 80 groups of parameter combinations mentioned above. The first and second rows list the original input images and corresponding ground truth images, respectively. The third and fourth rows show the best results for the contour detectors with anisotropic and isotropic inhibition of Grigorescu et al. (2003), respectively. The last row shows the results of our contour model. Table 1 quantitatively lists the model parameters and performance measures, as well as the percentages of false positives and false negatives, for the four images shown in Fig. 5.

The results clearly show that our computational model suppresses the edges originating from textural regions like foliage or grass much more effectively than the other two inhibition-based models, which is graphically reflected by the much fewer trivial edge fragments in the textural regions of the images shown in Fig. 5 and, meanwhile, is quantitatively revealed by the significantly smaller measure e_{FP} listed in Table 1. In addition, our model improves the continuity of object contour to some extent, as indicated by the relatively lower measure e_{FN} in Table 1. As a typical example, our model produces a result that is quite close to the ground truth for the Hyena image shown in the first column of Fig. 5.

For each of the 40 test images, the 80 different P values of each model computed with the 80 groups of different parameter combinations specified earlier are statistically illustrated in Fig. 6 by using the statistical box-and-whisker plots. Graphically, the top end of a whisker corresponds to the best performance of a specific case, and the horizontal red line in the box of a whisker shows the median value of performance measure in this study. It should be clear from the plots that our model achieves a consistent better performance over the previous anisotropic and isotropic inhibition models. Specifically, the whisker's top ends of our model are notably higher than those of other two models for most images. In addition, our model produces a significantly higher median performance measure in almost all cases and meaningfully smaller range between the lower and upper quartiles for most of the cases, revealing that with a more robust manner, our model performs better to extract salient contours from cluttered scenes.

Note that the presented results of anisotropic and isotropic inhibition models shown in Fig. 5 and Table 1 are not exactly identical to those listed in the work of Grigorescu et al. (2003). This slight difference is mainly, perhaps, due to the difference in implementation details between the two works. This difference, however, does not affect our performance comparison and the deduced conclusions, as different models in this study remain identical values for those common parameters.

Summary and discussion

Summary

In this paper, we proposed an adaptive inhibition-based contour detection model that outperforms standard inhibition-based detectors when extracting object contours in cluttered natural scene images. Specifically, the model we have studied and implemented is strongly motivated by biological evidences as well as computational considerations. First, inspired by the recent findings of physiological experiments and psychophysical studies about the center–surround interaction in primary visual cortex of biological vision system (Jones et al., 2001; Kapadia et al., 2000; Li, 1996; Li and Li, 1994, 1995; Walker et al., 1999, 2000; Yao and Li, 1998), we introduced into our model context-sensitive inhibition from end regions of nCRF while keeping consistent side inhibition irrespective of contextual information. Namely, the inhibition coming from end regions decreases adaptively with increasing possibility of object contour occurring, whereas the end inhibition increases to be as strong as side inhibition when the local image patch is most likely to be filled with meaningless textures or cluttered background. Compared with those previous inhibition models (Grigorescu et al., 2003; Papari et al., 2007; Petkov and Westenberg, 2003), such adaption helps our model avoid the problem of self-inhibition with a more biologically plausible scheme and, consequently, leads to a more effective suppression of textures and at the same time a significant improvement in protecting object contours being destroyed.

The second characteristic of vision system introduced into our model is the multi-scale or multi-resolution information processing mechanism. In fact, there has been a growing interest in the use of multi-scale analysis for image processing such as contour detection (Joshi and Sivaswamy, 2005; La Cara and Ursino, 2008; Liang et al., 1999; Lindeberg, 1998; Mallat, 1989; Papari et al., 2007; Ren, 2008), based on extensive experimental evidence supporting that multi-scale mechanism is psychophysically plausible for image perception (Campbell and Robson, 1968; Julesz, 1962; Mermillod et al., 2005; Morrone and Burr, 1997; Oliva and Schyns, 1997). However, the role of interactions that exist between different scales in biological vision system is not well understood. Computationally, how to combine the information at different scales is challenging because of the inaccurate localization of edges at coarser scales and the overfull details at finer scales (Liang et al., 1999; Papari et al., 2007). In this study, we proposed a simple but effective way to combine the information only from two scales. The coarse-scale information in our model works as a kind of *saliency map* (Li, 2002) to identify where objects are. Instead for localizing, this orientation based saliency map is employed as a measurement for lessening end suppression of possible contours. In contrast, the accurate extraction and localization of object contours is realized at the fine scale. Such two-scale based model provides an innovative combination of different scales information to realize the adaptive end inhibition while subtly avoiding the problems occurring at a single scale.

Biological plausibility and a possible neural network

Increasing studies on the response properties of neurons in primary visual cortex (V1) suggest strongly that V1 neurons act as an adaptive information processor whose processing is profoundly influenced by visual context, top–down control, and training (Gilbert and Sigman, 2007; Li et al., 2006), and this adaptability originates from the dynamic and complex nature of the relationship between the classical receptive field and surrounding regions of visual space (Bredfeldt and Ringach, 2002; Dragoi and Sur, 2000; Gilbert, 1992; Gilbert et al., 1996; Kapadia et al., 1999; Shapley et al., 2003; Walker et al., 1999; see DeAngelis et al., 1995; Fitzpatrick, 2000 for review). In the following parts, the general biological plausibilities of the

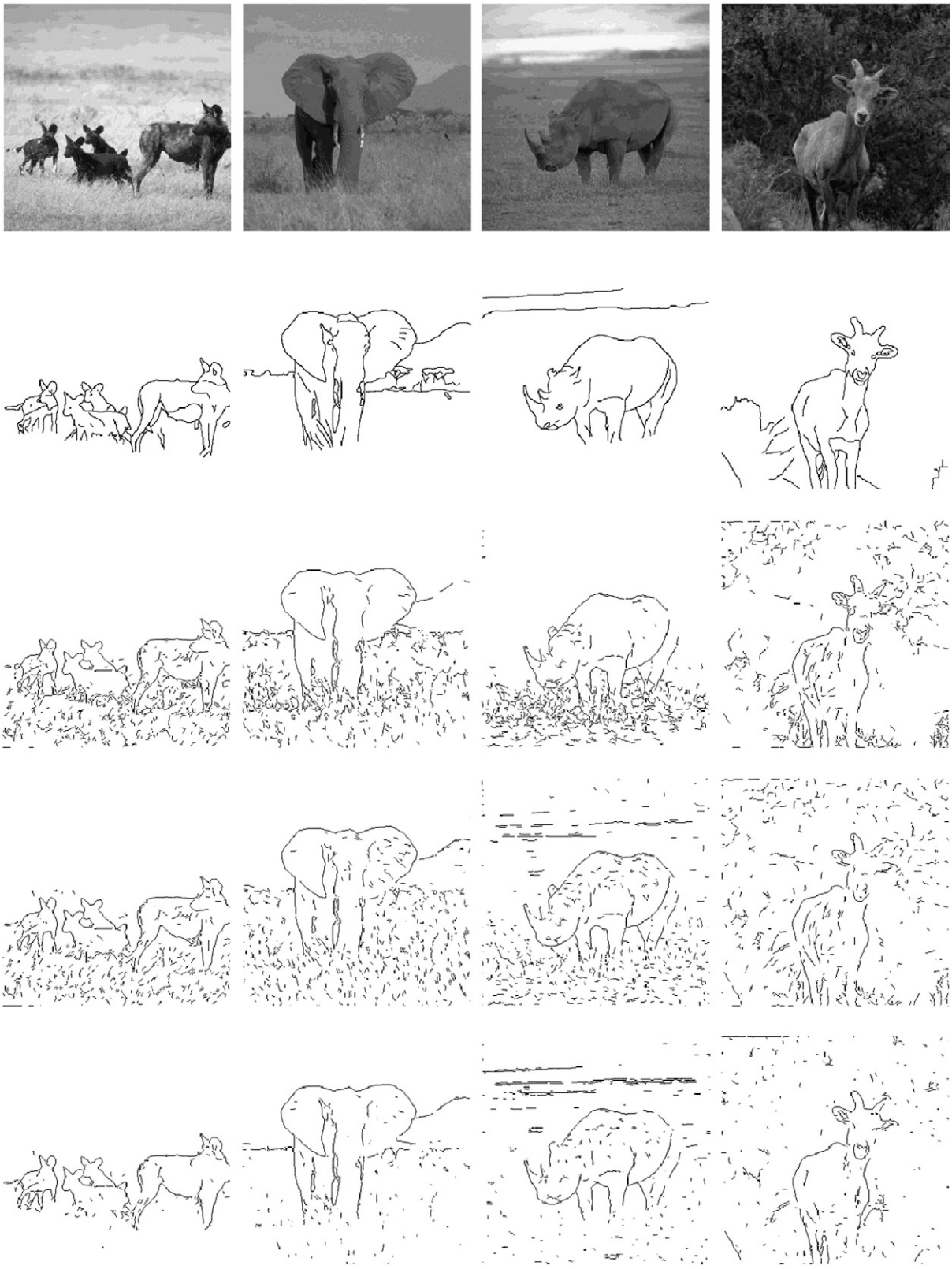


Fig. 5. The best results of contour detection on four natural images. The rows from top to bottom represent the original image, ground-truth contour, output of anisotropic and isotropic inhibition detectors (Grigorescu et al., 2003) and our model, respectively.

proposed model are first discussed, followed by a possible neural network underlying the proposed model.

One of the general biological plausibilities of the proposed model comes from the sufficient physiological evidence showing the spatially localized and functionally distinct components of CRF surrounds of neurons in primary visual cortex (Cavanaugh et al., 2002; DeAngelis et al., 1994; Freeman et al., 2001; Jones et al., 2001; Kapadia et al., 2000; Li and Li, 1994; Maffei and Fiorentini, 1976; Sceniak et al., 1999, 2001; Walker et al., 1999, 2002). Li and Li (1994) and DeAngelis et al. (1994) performed systematic studies to

Table 1
Model parameters and performances for the images shown in Fig. 5.

Image	Model	σ_f	σ_c	p	α_1	α_2	e_{FP}	e_{FN}	P
Hyena	Our model	2.00	10.00	0.10	1.00	1.00	0.23	0.17	0.70
	Anisotropic	2.40		0.10	1.20		0.80	0.20	0.49
	Isotropic	2.40		0.10	1.00		0.73	0.19	0.51
Elephant 2	Our model	2.00	12.00	0.10	1.00	1.00	0.36	0.27	0.57
	Anisotropic	2.40		0.10	1.20		1.17	0.27	0.39
	Isotropic	2.40		0.10	1.20		0.98	0.27	0.42
Rino	Our model	2.00	10.00	0.10	1.00	1.20	0.54	0.32	0.39
	Anisotropic	2.40		0.10	1.00		2.13	0.37	0.27
	Isotropic	2.40		0.10	1.20		1.65	0.28	0.33
Goat 3	Our model	2.00	10.00	0.10	1.00	1.20	0.91	0.31	0.42
	Anisotropic	2.20		0.10	1.00		1.47	0.26	0.36
	Isotropic	2.40		0.10	1.00		1.36	0.33	0.35

quantitatively characterize the spatial properties, especially the relationship between end and side inhibition, of cat's striate cortical neurons. Specifically, Li and Li (1994) found that width and length modulation (facilitation or inhibition) for a given cell could be similar or different, resulting in four possible combinations for surround organization. It is suggested that the different surround structures may be adapted to analyze different texture patterns of visual images (Li, 1996). Subsequently, spatially asymmetric surround suppression was observed by many investigators (Cavanaugh et al., 2002; Freeman et al., 2001; Jones et al., 2001; Sceniak et al., 2001; Walker et al., 1999), who found that suppression may originate from any specific localized region of the surround. It is suggested that localized surround effects may be a substrate for figure–ground segmentation of visual scenes (Walker et al., 1999, 2002). The functional significance of such spatial organization of surrounds is still not quite clear; however, it is at least indicated that any computational model that tries to utilize surround inhibition should account for a wide variety of spatial organizations (Fitzpatrick, 2000; Li and Li, 1994; Walker et al., 1999, 2002). From the viewpoint of computational modeling, surround modulation effects are asymmetrically weighted around the RF center (Walker et al., 2000).

Another general biological plausibility of the present model comes from the neurophysiological findings of Kapadia et al. (1999, 2000) about the contrast-dependent surround modulation. At low contrast (20%), strong excitatory end interactions were evident along the collinear axis, while at relatively higher contrast (50%), collinear facilitation along end regions was quite weak. Meanwhile, the inhibition was always strong at the lateral position (i.e., side regions). Similar observations were reached by Sceniak et al. (1999). Although the existence of facilitation has still remained a matter of controversy mainly because of the difference in determining the extent of the central excitatory receptive field and the stimuli patterns adopted in different studies (see Angelucci and Bressloff, 2006; Fitzpatrick, 2000 for review), these findings show a clear common trend, that is, (1) for some cells in primary visual cortex, the presenting of stimuli within the side regions always induces a strong inhibition effect on the cell's response at medium-to-high contrasts, which is computationally simulated as the consistent side inhibition in the present model. (2) In addition, end inhibition is contrast-dependent (see discussion in Kapadia et al., 2000), that is, the presenting of collinear stimuli at low-to-high contrasts in the end regions may increase the cell's response, compared to all other cases, no matter the response increases to be higher (Kapadia et al., 1995, 1999, 2000; Li and Li, 1994; Sceniak et al., 1999) or still lower (DeAngelis et al., 1994; Walker et al., 1999) than the response to the stimuli presented in the CRF alone. In the present model, such trend of increasing of cell's response to collinear stimuli within end regions is computationally modeled by decreasing the end inhibition with the increasing of Gabor energy at a coarser scale.

As for the possible biological and anatomical substrates underlying the center–surround interactions, most studies suggest that these interactions are primarily mediated by the network of long-range horizontal intracortical connections that originate from excitatory cortical pyramidal cells of V1 (Gilbert and Wiesel, 1989; Ts'o et al., 1986; see Gilbert, 1992; Fitzpatrick, 2000; Series et al., 2003 for review). In short, these intrinsic connections are widespread and may reach both excitatory and inhibitory post-synaptic cells, generating about 80% of long-range excitatory connections and 20% inhibitory connections (Hirsch and Gilbert, 1991; McGuire et al., 1991). Beside the horizontal intracortical connections, more and more physiological and psychophysical results indicate that feedback projections from higher cortical areas (e.g., V2, V4) and subcortical mechanisms, e.g., the length tuning properties of LGN neurons (see Fitzpatrick, 2000; Series et al., 2003; Angelucci and Bressloff, 2006 for review), also provide clear substrate for center–surround interactions in V1.

As for the present model, the involved mechanisms could be explained by several kinds of intracortical connections among V1 neurons and feedback connections from extrastriate cortex, and the basic architecture of a possible neural network with these mechanisms is schematically illustrated in Fig. 7. In particular, three types of inhibition are involved: (1) long-range intracortical isotropic inhibition coming from both side and end regions, (2) decrease of excitatory inputs via long-range intracortical inhibition contributed by the simultaneous stimulating of side and end regions, (3) decrease of end inhibition via feedback projection from extrastriate cortex. They are explained in details as follows.

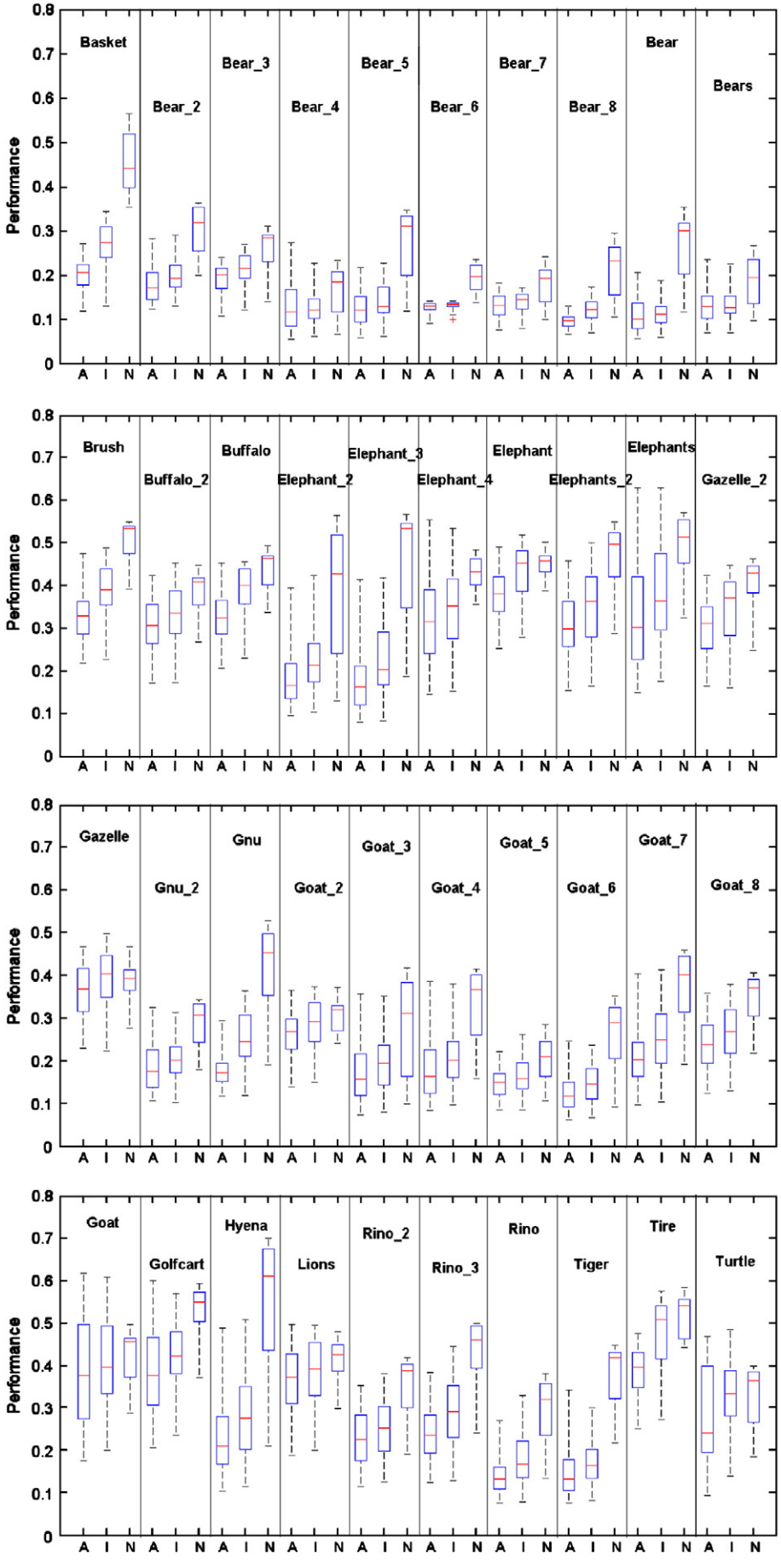
Long-range intracortical isotropic inhibition (coming from both side and end regions)

There is sufficient evidence suggesting that, besides the stronger long-range projections to iso-oriented pyramidal cells and nearby inhibitory cells in V1, there also exist weaker projections to non-iso-oriented pyramidal cells and neighboring inhibitory cells (Dragoi and Sur, 2000; Gilbert and Wiesel, 1989; Weliky et al., 1995). Specifically, a large weaker field of diffuse isotropic suppression was observed around the RF (Kapadia et al., 2000). Furthermore, at the single-cell level, several studies found that about 20–30% primate V1 cells show nonspecific inhibitory effects (Jones et al., 2001; Knierim and Van Essen, 1992; Nothdurft et al., 1999), i.e., the responses of the so-called generally suppressed cells to the lines in central RF are greatly suppressed by both optimal and orthogonal texture surrounds.

We think that for the specific task of contour detection under real-world conditions, where objects are most likely to be embedded in complex scenes with randomly oriented background (see the images shown in Fig. 5 as typical examples), general suppression with isotropic effects may play a critical role to make the object contours pop out by greatly suppressing the cluttered background, although iso-orientation facilitation and inhibition may play same important roles for most perceptual tasks (Kapadia et al., 2000). In fact, the mechanism of general suppression has been computationally demonstrated that it serves to suppress the randomly oriented textures effectively (Grigorescu et al., 2003; Petkov and Westenberg, 2003; Ursino and La Cara, 2004).

Taken together, we only consider the isotropic surround inhibition in this study and explain its neural substrates using long-range horizontal connections, as suggested in the model of Ursino and La Cara (2004). Considering that in our computational model, the subsequent surround inhibitions are computed on the basis of Gabor energy (see Eq. (3)), which simulates the neuronal responses elicited by the stimuli within RF center alone, here we ignore the feed-forward inhibition generated by monosynaptic thalamocortical input (see further discussion later). These explanations hold for both the

Fig. 6. Box-and-whisker plots of the performance of anisotropic inhibition model (denoted by A), isotropic inhibition model (denoted by I), and our contour model (denoted by N) for the 40 test images used in this study.


Basket
Bear_2
Bear_3
Bear_4
Bear_5
Bear_6
Bear_7
Bear_8
Bear
Bears
Brush
Buffalo_2
Buffalo
Elephant_2
Elephant_3
Elephant_4
Elephant
Elephants_2
Elephants
Gazelle_2
Gazelle
Gnu_2
Gnu
Goat_2
Goat_3
Goat_4
Goat_5
Goat_6
Goat_7
Goat_8
Goat
Golfcart
Hyena
Lions
Rino_2
Rino_3
Rino
Tiger
Tire
Turtle
Performance
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
A I N

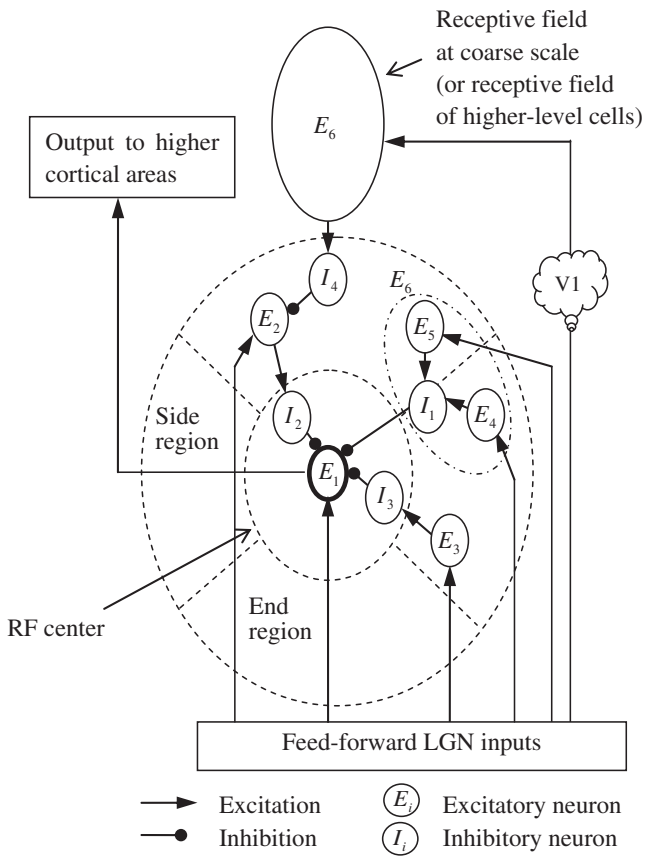


Fig. 7. Schematic drawings illustrating a possible neural network with the involved mechanisms in the present model. For the graphical clarity, the cells are drawn almost separately in the spatial space, and only one side region and one end region are illustrated in details. Excitatory neuron E_1 is identified as the target neuron in the receptive field (RF) center whose response we wish to determine. We assume that the preferred orientation of the target neuron E_1 is vertical, so in its surrounding area, the two end regions are vertically aligned, and the two side regions horizontally aligned. The target neuron receives input thalamocortical input via the direct feed-forward connections and surrounds modulations via lateral connections. Note that the inhibitory interneurons I_1, I_2, and I_3 are located in the same or nearby hypercolumns to that of target neuron E_1, and the current drawing is not at true scale for the sake of graphical clarity. The involved three surround inhibition mechanisms are summarized as follows. (1) Long-range intracortical isotropic inhibition coming from both side and end region: LGN→E_3→I_3→E_1 for side inhibition and LGN→E_2→I_2→E_1 for end inhibition. (2) Decrease of excitatory inputs via long-range intracortical inhibition: when both side regions and end regions are stimulated simultaneously, the activity of the length-tuned neuron E_6 in layer 4 may be reduced, resulting in a decrease of excitatory input to E_1 (here the activity decrease of the length-tuned neuron E_6 is graphically realized by a combination of E_4, E_5, and I_1). (3) Feedback projection from extrastriate cortex: the low-frequency LGN inputs reach fast to V1 and then higher level area, which feedback to V1 by targeting I_4, and then modulate E_1 via E_2 and I_2.

consistent side inhibition given by Eq. (10) and the first component of adaptive end inhibition given by Eq. (16).

As shown in Fig. 7, besides receiving direct feed-forward excitatory input from the lateral geniculate nucleus (LGN), the target neuron E_1 in layer 4 also receives feed-forward inhibition from cortical inhibitory interneurons I_2 and I_3 located in the same hypercolumn or in close hypercolumns to that of E_1. That is, the activated neuron E_2 in the end regions activates the inhibitory interneuron I_2 via long-range connection, which in turn inhibits the target neuron E_1 via local connection. Similarly, side inhibition works via the routine as LGN→E_3→I_3→E_1.

The mechanisms of long-range intracortical connections have been computationally simulated in many neural network models (Dragoi and Sur, 2000; Hansen and Neumann, 2008; Li, 1998; Pettet et al., 1998; Ross et al., 2000; Somers et al., 1998; Ursino and La Cara, 2004; Yen and Finkel, 1998). Specifically, in the neural network model of Ursino and La Cara (2004), besides the local inhibition via short-range feed-forward connections, the inhibitory modulation from far surround is simulated by a diffuse inhibition via a kind of long-range feedback intracortical connections. In the computational implementation of our model, the isotropic inhibition is computed in a region almost covering both the near and far surround (see the distance related weighting templates given by Eqs. (11) and (14)), for both the side regions (Eq. (10)) and end regions (Eqs. (15) and (16)). Recent studies suggested that horizontal connections from far surround to RF center are too slow, and conversely, the fast-conducting extrastriate feedback connections to V1 may be the possible substrate of far-surround modulation (Angelucci and Bressloff, 2006; Angelucci and Bullier, 2003; Bair et al., 2003; Ichida et al., 2007). To combine the large weaker field of diffuse isotropic suppression (Kapadia et al., 2000) in a more biologically plausible manner, an improved version of the proposed possible neural network model shown in Fig. 7 is to introduce extrastriate feedback connections from far surround to the central target neuron, i.e., thalamocortical inputs activate excitatory neurons located in far surround of side and end regions, which then activate neurons in extrastriate areas. The activated extrastriate neurons send feedback projections to excitatory V1 neurons in the RF center or near surround, which in turn activate the nearby inhibitory interneurons and then inhibit the target cell (Ichida et al., 2007).

Decrease of excitatory inputs via long-range intracortical inhibition (contributed by the simultaneous stimulating of side and end regions)

Besides the physiological ratio of ~80% excitatory to ~20% inhibitory neurons in V1 (Hirsch and Gilbert, 1991; McGuire et al., 1991), a predominant fraction of excitatory long-range connections was also quantitatively observed. A recent work of Stepanyants et al. (2009) shows that almost 74% of excitatory inputs to neurons arises from outside their resident hypercolumn via long-range lateral connections. Such an unexpectedly large fraction of excitatory inputs to the target neuron originating from distant surrounding neurons indicates that the stimuli presented in surrounding areas may influence directly the total excitatory inputs to the studied target cell (Stepanyants et al., 2009).

It has been generally accepted that the excitatory inputs to a V1 pyramidal neuron may arise from (see Angelucci et al., 2002 for review) (1) feed-forward afferents from the thalamus and other V1 layers; (2) long-range lateral connections from distant pyramidal neurons; and (3) feedback connections from extrastriate cortex. To investigate the roles of excitatory inputs, as well as inhibitory inputs, on the property of length tuning, Anderson et al. (2001) measured length tuning of cells in cat primary visual cortex with intracellular recordings. With the aid of a current-injection technique to separate excitatory and inhibitory influences, they found that, compared with optimal stimuli, excitatory synaptic conductance generally decreases with stimulus length, and inhibitory conductance generally shows averaged decrease, normally with two peaks, one for short stimuli and one for stimuli with intermediate length. Recently, Ozeki et al. (2009) observed the similar phenomena showing that increases in visual stimulation beyond the cat V1 neuron's optimal size caused decreases in both excitatory and inhibitory conductance.

As for the drop in excitatory conductance, a possible explanation is the contribution of subcortical mechanism, i.e., the excitatory responses of LGN cells may be suppressed by stimuli outside their receptive fields (Anderson et al., 2001; Ozeki et al., 2004). Another possibility, however, is more likely that cortical excitatory inputs from other cortical cells that are themselves length-tuned resulted in a decrease in excitation with stimulus length (Anderson et al., 2001; Ozeki et al., 2009). To relate our model to this biological possibility, we postulate that, when strong stimuli cover both the end and side regions of V1 cells, there is a higher chance for those length-tuned cells around the studied cell to be suppressed. A possible manner is

that those length-tuned layer 6 cells may reduce the excitatory input to layer 4 cells, since besides some layer 6 cells with long receptive fields (Bolz and Gilbert, 1986; Bolz et al., 1989), it has been also reported that a large proportion of layer 6 cells of cat's primary visual cortex have short receptive fields, some of which are themselves length tuned (DeAngelis et al., 1994; Grieve and Sillito, 1991a). In addition, there is good evidence to support the existence of strong excitatory influence on visual responses of hypercomplex cells from layer 6 to layer 4 (Ferster and Lindstrom, 1985; Grieve and Sillito, 1991b). Taken together, layer 4 cells may receive length-tuned excitatory inputs from layer 6 cells (Grieve and Sillito, 1991b).

Without excluding all the possible mechanisms underlying the decrease in excitatory inputs with stimulus length mentioned above, we use a generalized neural circuitry to simulate this phenomenon. For clarity, the decrease in excitatory inputs to target neuron E_1 is graphically realized in Fig. 7 by a combination of E_4, E_5, and I_1. For example, the combination of E_4, E_5, and I_1 may be regarded as a length-tuned neuron E_6 in layer 4. When both side regions (E_4) and end regions (E_5) are stimulated simultaneously, the excitatory inputs to the interneuron I_1 may be large enough to activate I_1, and I_1 in turn inhibits the studied neuron E_1, resulting in a decrease of excitatory input to E_1.

Since intrinsic cortical connections are widespread (Gilbert and Wiesel, 1989), and besides the stronger long-range projections to iso-oriented pyramidal cells and nearby inhibitory cells in V1, cortical cells also have weaker projections to non-iso-oriented pyramidal cells and neighboring inhibitory cells (Dragoi and Sur, 2000; Gilbert and Wiesel, 1989; Weliky et al., 1995), it is reasonable to presume that there are strong or weak connections between the pyramidal cells within the surrounding regions (Angelucci et al., 2002). Furthermore, there is also evidence indicating that side and end inhibition may have distinct underlying neural substrates. Li and Li (1994) and DeAngelis et al. (1994) found that for some cat's striate cortical cells, the inhibition effect is very weak for the side regions and very strong for the end regions. Subsequently, Walker et al. (1999) and Freeman et al. (2001) observed that although the maximal suppression can arise from any location of the surrounding areas, there is a clear bias for maximal suppression to arise from the end zones. Sceniak et al. (2001) observed the similar trend for some V1 neurons of macaque monkey that suppression strength along the end axis tends to be significantly greater than side suppression. The functional implication of asymmetrical surround suppression is still unclear (Fitzpatrick, 2000; Series et al., 2003). It at least, however, reasonably suggests that the side regions and end regions may have different underlying cortical circuitry (Li and Li, 1994; Yazdanbakhsh and Livingstone, 2006). In addition, physiological findings have demonstrated a tendency that end-inhibited neurons also show side suppression (DeAngelis et al., 1994; Li and Li, 1994). Such positive correlation between end and side inhibition suggests that these two phenomena may be inherently related (Skottun, 1998). In fact, Knierim and van Essen (1992) have suggested the possible existence of some interesting interactions between end regions and side bands beyond the CRF, and Series et al. (2003) have proposed the subtle interplay between a variety of circuits that are simultaneously activated. Further detailed studies on the interactions among the neurons in the surround of the studied target cell may help reveal exact underlying mechanisms of those interesting observations on the center–surround interactions of the studied cell (Fitzpatrick, 2000; Ohki and Reid, 2007; Walker et al., 2002).

From the viewpoint of computational modeling, the possible interaction between side and end regions mentioned above is consistent with the nonlinear property of center–surround interaction. It has been demonstrated by quantitative comparisons that the actual inhibition of the entire surround is not simply a linear sum of the inhibition of the side and end regions, and the entire surround inhibition may be larger, equal, or less than the linear sum of side and end inhibitions (Kastner et al., 1999; Knierim and Van Essen, 1992; Li and Li, 1994).Without ruling out other possibilities, we speculate that the interactions between the side regions and end regions may partially contribute to the nonlinear property of center–surround interaction.

Feedback projection from extrastriate cortex

The direct biological evidence for the introduced mechanism that information at coarse scale adjusts adaptively the V1 center–surround interaction comes from the following physiological findings.

(i) Some neurons in higher level areas are activated sufficiently rapidly to be able to have enough time to influence the neurons in lower level areas such as V1 (Girard et al., 2001; see Bullier, 2001; Kveraga et al., 2007 for review). Specifically, the activity transferred through magnocellular cells (M cells) of the LGN may reach area V1 much earlier than activity transferred from the parvocellular cells (P cells) of the LGN, and thus, M activity generally precedes P activity in the different layers of V1 and the higher level areas along the dorsal stream (i.e., the "where" or parietal stream). Most likely, the early activation of the areas along the dorsal pathway is attributed to its almost exclusive drive by the so-called M channel and the high degree of myelination in many dorsal areas, such as MT (medial temporal area), MST (medial superior temporal area), and FEF (frontal eye field) (Bullier, 2001; Callaway, 1998; Maunsell et al., 1990). It is worth noting that the dorsal visual pathway is generally believed to specialize in information processing about spatial location and motion. However, there is increasing evidence showing that areas in the dorsal pathway also contribute to represent shape information and visual category information, perhaps at a quite coarse-grained level (Freedman and Assad, 2006; Lehky and Sereno, 2007; see Hegde, 2008 for review). Either coarse spatial locations or object shapes are useful information for the specific task of fine contour detection in the present model.

(ii) The feedback connections from higher level areas appear sufficiently rapid to reach lower level areas such as V1 in a few milliseconds (Girard et al., 2001; Kveraga et al., 2007; Movshon and Newsome, 1996). A substantial evidence is that the conduction velocities of feed-forward and feedback connections between areas V1 and V2 of the monkey are almost ten times faster than that of horizontal intracortical connections in V1 (around 3.5 m/s vs. 0.33 m/s) (Girard et al., 2001; see Angelucci and Bullier, 2003 for review). In addition, some physiological findings suggest that feedback from MT may act on V1 neurons almost at the same time or even before information from the LGN reaches these cells (see Lamme et al., 1998 for review). Such a rapid conduction of feedback connections provides a chance for higher level areas to influence in time the responses of V1 neurons to the later arrival of feed-forward input through the P cells in LGN, together with the slow surround modulation through long-range horizontal intracortical connections in V1 (Bullier, 2001; Kveraga et al., 2007).

(iii) It is most likely the low spatial frequencies (LFs) in the image that are projected rapidly by anatomical shortcuts from early visual areas directly to the higher level areas, possible along the dorsal stream (Bar et al., 2006; see Hegde, 2008; Kveraga et al., 2007 for review). The subcortical shortcuts include the Y retinal ganglion cells responding to low-resolution stimuli and the magnocellular cells in LGN receiving preferentially projections from Y retinal ganglion cells. Thereafter, that difference between low- and high-spatial frequencies still remains at the level of the cortical cells. In particular, LFs in the image are

most likely projected rapidly from V1 to the higher dorsal areas via the magnocellular pathway (Bar et al., 2006; Hegde, 2008).

Taken together, both anatomical and neurophysiological findings suggest that the neural circuit between V1/V2 and MT is most likely to subserve rapid top–down feedback for center–surround interaction and figure–ground segregation in primary visual cortex (Lamme et al., 1998; see Likova and Tyler, 2008 for discussion). Considering the observation that top–down effects are most strongly exerted on contextual influences, Gilbert and Sigman (2007) suggested that perceptual learning involved in the tasks of shape discrimination and contour integration needs appropriate interactions between feedback and the local circuits that provide contextual information, and there are particular ensembles of feedback connections serving to selectively modulate subsets of horizontal connections. On the basis of this idea, Gilbert and Sigman (2007) proposed a top–down modulation model, in which feedback connections target inhibitory interneurons in V1 so as to neutralize the unwanted suppressive activity of intracortical horizontal connections within V1 when objective contour occurs.

Similar to the proposal of Gilbert and Sigman (2007), here we proposed a possible neural circuit for top–down modulation using the coarse information, as shown in Fig. 7. The global low-frequency features project fast to V1 and then to the higher level areas, resulting in a higher level neuron E_6 with longer RF to be activated. Via feedback connections, E_6 then projects to the inhibitory interneuron I_4 within the end regions of the target neuron E_1 in V1, and the activated I_4 inhibits the excitatory pyramidal cell E_2, which in effect reduces the strength of inhibition on the target neuron E_1 via the inhibitory interneuron I_2.

Functionally, the earlier processed global information characterized by low spatial frequency may serve to guide the subsequent processing (Hegde, 2008; Kveraga et al., 2007). For instance, they could be used for an initial coarse segmentation of the stimuli, to be later refined by the slower accumulation of HF information (Goffaux et al., 2010). In fact, feedback connections from higher level areas to V1 seem to play a very important role in center–surround interactions of V1 cells (Angelucci and Bullier, 2003; Angelucci and Bressloff, 2006; Gilbert and Sigman, 2007).

The implemented feedback modulation in this study, modeled by a oriented Gabor with coarser scale in the present model, has the computational role to enhance the responses of target neuron if iso-oriented, coaligned cells in a broader neighboring region are simultaneously activated, and this enhancement of target cell's response is realized in our model as an overall decreasing of end inhibition by counteracting the inhibitory effects in the end regions originating from other two sources in the present model. It should be noted that Field et al. (1993) suggested that neighboring neurons tuned to similar orientations are linked through a local “association field”, a region around the element where other similar elements group together and segregate from the background. To a certain extent, the coarser scale Gabor filters in our model may also act as “association fields” to help integrate local stimuli with similar orientations.

Feedback mechanisms have been involved in increasing computational models. In particular, Li (1998) introduced additional top–down feedback connections into her model, together with horizontal cortical connections in V1, to realize selective contour enhancement given that the higher level areas already know which segments belong to a contour and what feedback signals to send back to V1. Considering the low conduction of horizontal connections in V1, Angelucci and Bressloff (2006) and Ichida et al. (2007) explicitly introduced extrastriate feedback connections from V2 to V1 into their neural network models as the possible substrate of far-surround modulation. In a recent model of De Meyer and Spratling (2009), interactions between cortical top–down and horizontal connections in V1 were modeled to explain the attentional modulation of collinear facilitation in V1. Different with these models, the possible model shown in Fig. 7 proposes that it is the global information characterized by low spatial frequency that serves as a kind of fast feedback signals to modulate (enhance) the detection of fine contours by reducing the end inhibition if needed. Note that we do not rule out other possible feedback channels suggested in other successful models (Angelucci and Bressloff, 2006; De Meyer and Spratling, 2009; Ichida et al., 2007; Li, 1998). Here we just emphasize the possible roles of low spatial frequencies in the specific task of contour detection.

Recent experimental data from visual cognitive sciences suggest that visual information is integrated in a coarse-to-fine manner, i.e., visual perception initially uses coarse blobs of the image before entering into finer details (Menz and Freeman, 2003; Mermillod et al., 2005; Oliva and Schyns, 1997; Parker et al., 1997). Such coarse-to-fine mechanism indicates that it may be computationally beneficial to employ early low-frequency information to guide or constrain the subsequent fine analysis of information with higher frequencies in the visual field. However, many neural substrates for coarse-to-fine processing are far from clear due to the fact that the central visual system is so extremely complex that many higher level areas are poorly understood (Hegde, 2008). For example, extensive further investigations are needed to clarify the question that whether and to what extent magnocellular pathway transfer lower spatial frequencies faster than parvocellular pathway (Hegde, 2008; Kveraga et al., 2007).

Before leaving the explanations of biological plausibility of the present model, another point that should be pointed out is, in our proposed possible neural network model shown in Fig. 7, only intracortical connections and a kind of extrastriate feedback connections are used to account for the surround modulation mechanisms involved in our computational model. Note that there is also sufficient evidence showing that feed-forward afferents from LGN, which mainly target neurons in V1 layer 4C, contribute to V1 surrounds (Fitzpatrick, 2000). In feed-forward inhibitory circuits, the stimulation of an afferent fiber may directly excite an inhibitory cell, which in turn inhibits the firing of the target neuron. Such feed-forward inhibition mechanism was originally modeled in an explicit form by Ursino and La Cara (2004), who proposed that inhibitory interneurons located near the target cell are activated by thalamic input, and the activated interneurons in turn inhibit the target cell via short-range connections. Functionally, such feed-forward connections from inhibitory interneurons may help contour extraction by selectively inhibiting the response of cortical cells to non-optimally oriented stimuli (Petkov and Westenberg, 2003; Ursino and La Cara, 2004) or by isotropically suppressing the response to randomly oriented background (Nothdurft et al., 1999; Petkov and Westenberg, 2003; Ursino and La Cara, 2004).

Other related works

Increasing physiological studies (Fitzpatrick, 1996; Gilbert, 1992; Gilbert et al., 1996; Kapadia et al., 1995, 1999, 2000; Li et al., 2006), psychophysical investigations (Dakin and Baruch, 2009; Field et al., 1993; Kapadia et al., 1995, 2000; Kovacs and Julesz, 1993; Polat and Sagi, 1994), and computational modeling works (Hansen and Neumann, 2008; Li, 1998; Pettet et al., 1998; Ursino and La Cara, 2004; Yen and Finkel, 1998) suggest that, although contextual interaction and contour integration can extend across large areas of visual space, this spatial integration may occur even at the primary visual cortex (V1), the first stage of visual cortical processing, indicating that V1 plays an important role of stimulus-driven, bottom–up processes in pre-attentive contour integration. In particular, it has been long recognized that the possible contour integration in V1 may depend primarily on excitatory horizontal connections found predominantly in the superficial layers 2/3 of V1, which

preferentially link sets of neurons with similar functional properties, such as orientation preference (Field et al., 1993; Gilbert and Wiesel, 1989; Rockland and Lund, 1983; Ts'o et al., 1986), especially if their receptive fields are approximately displaced along an axis of preferred orientation in visual space (Li et al., 2006; Nelson and Frost, 1985; Kapadia et al., 1995, 1999, 2000; Shouval et al., 2000).

It has been well accepted that contour integration follows the Gestalt law of "good continuation" (Wertheimer, 1923; see Gilbert and Sigman, 2007 for a recent review). The findings mentioned above suggest that the Gestalt law of "good continuation" or "smoothness" that directs contour integration may have its clear neural substrate in V1. In terms of computational modeling, various works tried to implement these mechanisms underlying the Gestalt "good continuation" or "smoothness" rule (Dakin and Baruch, 2009; Li, 1998, 1999, 2002; Papari and Petkov, 2008; Pettet et al., 1998; Sigman et al., 2001; Ursino and La Cara, 2004; Yen and Finkel, 1998). Specifically, Yen and Finkel (1998) used a co-circularity constraint to help group contour components. Ursino and La Cara (2004) used a mid-range cortical feedback excitation to implement the co-axial and co-modularity criteria to emphasize smooth contours. Obviously, the successful implementations of these models indeed help greatly formulate the Gestalt rules in an appropriate mathematical manner, and it is expected that Gestalt rule inspired computational implementation of neural mechanisms may significantly improve the results of contour extraction from a noisy background.

To a certain extent, our model presented here employs the law of "good continuation" by using the coarse Gabor energy map to guide the surround inhibition, because the continuation of object contours is generally much better than that of cluttered background, and this better continuation is reflected by the higher Gabor energy at a coarser spatial scale by binding those successive line segments belonging to a common contour. Psychophysically, the associated operations at the two coarse and fine scales in our model could be restrictedly interpreted as the subsequent pre-attentive and attentive stages in image perception (La Cara and Ursino, 2008; Li, 2002; Zhaoping and Dayan, 2006). Gabor filters with larger scales analyze the visual field at a coarser level of spatial frequency and try to pre-attend and identify where objects are. Subsequently, Gabor filters with smaller scales analyze the visual field at a very high-spatial frequency and generates a group of fine contours under the center–surround interaction. This, to a certain extent, simulates the interactions between a dorsal "where" and ventral "what" visual stream although the interactions between these two streams and the mechanisms of selective attention are far complicated than have been thought (Deco and Rolls, 2004).

The physiological investigations and computational studies mentioned above clearly suggest that more information is necessary to work in an interactive manner for the computational modeling of neural mechanisms underlying contour extraction. In fact, this idea has gained much attention from many researchers. For instance, previous models, including the present model, used a multi-scale approach (Joshi and Sivaswamy, 2005; La Cara and Ursino, 2008; Papari et al., 2007; Rodrigues and du Buf, 2006), and others proposed a selective control on inhibition (Ito et al., 1998; Ito and Gilbert, 1999; La Cara and Ursino, 2008). All of these multi-scale approaches mentioned here employ different scales to extract various contour details and then combine these details while eliminating noise with certain kind of surround inhibition. In particular, La Cara and Ursino (2008) significantly improved their former mathematical model of contour extraction (Ursino and La Cara, 2004) by including the mechanisms of multi-scale decomposition and attention. To a certain extent, our model presented in this work is most close to the model of La Cara and Ursino (2008), because the two models share two concepts. First, both models use only two scales. Second, a flexible surround inhibition is realized in the two models. However, there are differences in the implementation details about these two aspects. The main difference is that the spatial attention in the model of La Cara and Ursino (2008) works computationally by selecting an appropriate scale (spatial frequency channel) and the portion of the image to be scrutinized and modulating the strength of long-range inhibition by selecting the appropriate parameters about the strength of inhibition and its spatial extension. The circuits devoted to analysis of the image at two different scales are treated separately as two independent parallel circuits. Dissimilarly, our model includes a direct one-way interaction between the two scales, through which the information of coarser scale provides an adaptive modulation to the surround inhibition at the finer scale. Such one-way interaction is of course quite simplified, although working well for the test cases in this study. As pointed out by La Cara and Ursino (2008), although sufficient pieces of evidence from previous studies support the existence of parallel processing at different spatial scales, meticulous investigations are still required on the exact underlying mechanisms that regulate the possible interactions among them.

Discussion and conclusions

Computationally, a contour detector with consistent inhibition from whole nCRF or consistent facilitation from the end regions of nCRF needs to make a tradeoff between texture suppressing and contour retaining. In contrast, our proposed model improves the overall performance of contour detection mainly by suppressing more strongly the textures without sacrificing the continuity of object contour, which is computationally achieved by adaptively weighting the surround inhibition in different subregions. It could be optimistically expected that contour continuity would also be meaningfully improved when a mechanism of adaptive end facilitation is reasonably introduced into our model, which is our ongoing work.

As mentioned earlier, four possible combinations of side and end modulation have been supposed for the spatial organization of surround area (Li, 1996; Li and Li, 1994, 1995; Yao and Li, 1998). Since inhibitory modulations are more frequent and more pronounced (Li and Li, 1994; Walker et al., 2000), in this study, we focus on simulating the specific type A (*end*−/*side*−), but in an adaptive fashion. In general, most of the existing nCRF-based contour models employ the mechanism of surround inhibition. But now increasing physiological findings indicate that excitatory interactions may play an equally important role in neural responses (Kapadia et al., 1995, 1999, 2000; Li, 1999; Nelson and Frost, 1985; Polat et al., 1998; Sillito et al., 1995). Based on this spatial arrangement of contextual interactions, Tang et al. (2007) proposed a novel contour extraction model, in which a surround diagram type C (*end*+/*side*−) was employed to combine the dual processes of spatial facilitation and surround inhibition for contour extraction from complex scenes. End facilitation can indeed help find weak contours; however, it may also produce unwanted fake contours. As discussed by Tang et al. (2007), a flexible spatial facilitation is necessary.

Another point that deserves a brief comment is whether the present results suggest physiological experiments to validate or reject the underlying ideas. Although some aspects of our model are speculative, there are many pieces of evidence in the physiological literature supporting our most basic choices. Our model indicates that end inhibition is reduced when an object contour is being attended, which suggests that additional physiological experiments should focus on analysis of the strength of inhibition coming from different parts, especially from the end regions, of nCRF of V1 neurons at different levels of attentions with different parts of natural images containing objects as visual stimuli.

In conclusion, our proposed model simulates the behavior of V1 neurons in a manner of adaptive end inhibition. To our knowledge, this work is perhaps the first attempt in the literature to develop a

contour detector by introducing an adaptive center–surround interaction. The present model is built on the basis of Gabor-based phenomenological framework, but to a certain extent, the involved mechanisms could be biologically plausibly explained using the underlying neural circuits (although simplified) of vision process in V1. Besides the superior performance in contour detection, our model provides a better understanding on the roles of nCRF and has potential applications in computer vision and pattern recognition. In terms of contour detection, our ultra goal is to develop a more realistic model in which end regions could exert selectively modulatory influences of *inhibition* or *facilitation* with *adaptive* strength based on the local contextual features of complex scene. However, the precise adaptive relationship between the response of a neuron and its surrounding pattern when perceiving natural scenes is an extremely challenging issue that merits further investigation from both the computational and physiological aspects.

Acknowledgments

This work was supported by the Natural Science Foundations of China (90820301, 60835005, 30730036, and 60736029) and the Major State Basic Research Program of China (2007CB311001). This work was also supported by the Program for New Century Excellent Talents in the University of China (NCET-07-0151). The authors would like to thank Grigorescu, Petkov, and Westenberg for their natural image database with associated ground truth contour maps. We also thank the two anonymous reviewers for their valuable comments and suggestions on the manuscript.

References

Albright, T.D., Stoner, G.R., 2002. Contextual influences on visual processing. Annu. Rev. Neurosci. 25, 339–379.
Allman, J., Miezin, F., McGuinness, E., 1985. Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local–global comparisons in visual neurons. Annu. Rev. Neurosci. 8, 407–430.
Anderson, J.S., Lampl, I., Gillespie, D.C., Ferster, D., 2001. Membrane potential and conductance changes underlying length tuning of cells in cat primary visual cortex. J. Neurosci. 21, 2104–2112.
Angelucci, A., Bressloff, P.C., 2006. Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extraclassical receptive field surround of primate V1 neurons. Prog. Brain Res. 154, 93–120.
Angelucci, A., Bullier, J., 2003. Reaching beyond the classical receptive field of V1 neurons: horizontal or feedback axons? J. Physiol. Paris 97, 141–154.
Angelucci, A., Levitt, J.B., Lund, J.S., 2002. Anatomical origins of the classical receptive field and modulatory surround field of single neurons in macaque visual cortical area V1. Prog. Brain Res. 136, 373–388.
Bair, W., Cavanaugh, J.R., Movshon, J.A., 2003. Time course and time–distance relationships for surround suppression in macaque V1 neurons. J. Neurosci. 23, 7690–7701.
Bar, M., 2004. Visual objects in context. Nat. Rev. Neurosci. 5, 617–629.
Bar, M., Kassam, K.S., Ghuman, A.S., Boshyan, J., Schmid, A.M., Dale, A.M., Hamalainen, M.S., Halgren, E., 2006. Top–down facilitation of visual recognition. Proc. Natl Acad. Sci. USA 103, 449–454.
Bolz, J., Gilbert, C.D., 1986. Generation of end-inhibition in the visual cortex via interlaminar connections. Nature 320, 362–365.
Bolz, J., Gilbert, C.D., Wiesel, T.N., 1989. Pharmacological analysis of cortical circuitry. Trends Neurosci. 12, 292–296.
Bredfeldt, C.E., Ringach, D.L., 2002. Dynamics of spatial frequency tuning in macaque V1. J. Neurosci. 22, 1976–1984.
Bullier, J., 2001. Integrated model of visual processing. Brain Res. Rev. 36, 96–107.
Callaway, E.M., 1998. Local circuits in primary visual cortex of the macaque monkey. Annu. Rev. Neurosci. 21, 47–74.
Campbell, F., Robson, J., 1968. Application of Fourier analysis to the visibility of gratings. J. Physiol. 197, 551–566.
Canny, J., 1986. A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. 8 (6), 679–698.
Cavanaugh, J.R., Bair, W., Movshon, J.A., 2002. Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. J. Neurophysiol. 88 (5), 2547–2556.
Chan, W., Coghill, G., 2001. Text analysis using local energy. Pattern Recognit. 34, 2523–2532.
Chen, G., Dan, Y., Li, C.Y., 2005. Stimulation of non-classical receptive field enhances orientation selectivity in the cat. J. Physiol. 564, 233–243.
Dakin, S.C., Baruch, N.J., 2009. Context influences contour integration. J. Vis. 9 (2), 1–13.
Das, A., Gilbert, C.D., 1999. Topography of contextual modulations mediated by short-range interactions in primary visual cortex. Nature 399, 655–661.
Daugman, J.G., 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.
De Meyer, K., Spratling, M.W., 2009. A model of non-linear interactions between cortical top–down and horizontal connections explains the attentional gating of collinear facilitation. Vis. Res. 49, 533–568.
De Valois, R.L., Albrecht, D.G., Thorell, L.G., 1982. Spatial frequency selectivity of cells in macaque visual cortex. Vis. Res. 22, 545–559.
De Valois, K., Tootell, R., 1983. Spatial-frequency-specific inhibition in cat striate cortex cells. J. Physiol. Lond. 336, 359–376.
DeAngelis, G.C., Freeman, R.D., Ohzawa, I., 1994. Length and width tuning of neurons in the cat's primary visual cortex. J. Neurophysiol. 71, 347–374.
DeAngelis, G.C., Ohzawa, I., Freeman, R.D., 1995. Receptive-field dynamics in the central visual pathways. Trends Neurosci. 18 (10), 451–458.
Deco, G., Rolls, E.T., 2004. A neurodynamical cortical model of visual attention and invariant object recognition. Vis. Res. 44, 621–642.
Dobbins, A., Zucke, S.W., Cynader, M.S., 1987. Endstopped neurons in the visual cortex as a substrate for calculating curvature. Nature 329, 438–441.
Dragoi, V., Sur, M., 2000. Dynamic properties of recurrent inhibition in primary visual cortex: contrast and orientation dependence of contextual effects. J. Neurophysiol. 83, 1019–1030.
Ferster, D., Lindstrom, S., 1985. Synaptic excitation of neurones in area 17 of the cat by intracortical axon collaterals of cortico-geniculate cells. J. Physiol. Lond. 367, 233–252.
Field, D.J., Hayes, A., Hess, R.F., 1993. Contour integration by the human visual system: evidence for a local association field. Vis. Res. 33, 173–193.
Fitzpatrick, D., 1996. The functional organization of local circuits in visual cortex—insights from the study of tree shrew striate cortex. Cereb. Cortex 6, 329–341.
Fitzpatrick, D., 2000. Seeing beyond the receptive field in primary visual cortex. Curr. Opin. Neurobiol. 10, 438–443.
Forsyth, D., Ponce, J., 2003. Computer Vision: A Modern Approach. Prentice Hall.
Freedman, D.J., Assad, J.A., 2006. Experience-dependent representation of visual categories in parietal cortex. Nature 443, 85–88.
Freeman, R.D., Ohzawa, I., Walker, G.A., 2001. Beyond the classical receptive field in the visual cortex. Prog. Brain Res. 134, 157–170.
Gilbert, C.D., 1992. Horizontal integration and cortical dynamics. Neuron 9, 1–13.
Gilbert, C.D., Das, A., Ito, M., Kapadia, M.K., Westheimer, G., 1996. Spatial integration and cortical dynamics. Proc. Natl Acad. Sci. USA 93, 615–622.
Gilbert, C.D., Sigman, M., 2007. Brain states: top–down influences in sensory processing. Neuron 54, 677–696.
Gilbert, C.D., Wiesel, T.N., 1989. Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. J. Neurosci. 9 (7), 2432–2442.
Girard, P., Hupe, J.M., 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.
Goffaux, V., Peters, J., Haubrechts, J., Schiltz, C., Jansma, B., Goebel, R., 2010. From coarse to fine? Spatial and temporal dynamics of cortical face processing. Cereb. Cortex 21 (2), 467–476.
Grieve, K.L., Sillito, A.M., 1991a. Length summation properties of layer VI cells in the visual cortex and hypercomplex cell end zone inhibition. Exp. Brain Res. 84, 319–325.
Grieve, K.L., Sillito, A.M., 1991b. A re-appraisal of the role of layer VI of the visual cortex in the generation of cortical end inhibition. Exp. Brain Res. 87, 521–529.
Grigorescu, C., Petkov, N., Westenberg, M., 2003. Contour detection based on nonclassical receptive field inhibition. IEEE Trans. Image Process. 12 (10), 1274–1286.
Guo, K., Robertson, R.G., Mahmoodi, S., Young, M.P., 2005. Centre–surround interactions in response to natural scene stimulation in the primary visual cortex. Eur. J. Neurosci. 21, 536–548.
Hansen, T., Neumann, H., 2008. A recurrent model of contour integration in primary visual cortex. J. Vis. 8 (8), 1–25.
Haykin, S., 1998. Neural Networks: A Comprehensive Foundation2nd ed. Prentice Hall.
Hegde, J., 2008. Time course of visual perception: coarse-to-fine processing and beyond. Prog. Neurobiol. 84, 405–439.
Hirsch, J.A., Gilbert, C.D., 1991. Synaptic physiology of horizontal connections in the cat's visual cortex. J. Neurosci. 11, 1800–1809.
Hubel, D.H., Wiesel, T.N., 1968. Receptive fields and functional architecture of monkey striate cortex. J. Physiol. 195 (1), 215–243.
Ichida, J.M., Schwabe, L., Bressloff, P.C., Angelucci, A., 2007. Response facilitation from the "suppressive" receptive field surround of V1 neurons. J. Neurophysiol. 98, 2168–2181.
Ito, M., Gilbert, C.D., 1999. Attention modulates contextual influences in the primary visual cortex of alert monkeys. Neuron 22, 593–604.
Ito, M., Westheimer, G., Gilbert, C.D., 1998. Attention and perceptual learning modulate contextual influences on visual perception. Neuron 20, 1191–1197.
Jones, H.E., Grieve, K.L., Wang, W., Sillito, A.M., 2001. Surround suppression in primate V1. J. Neurophysiol. 86, 2011–2028.
Jones, J.P., Palmer, L.A., 1987. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. J. Neurophysiol. 58, 1233–1258.
Joshi, G.D., Sivaswamy, J., 2005. A Multiscale Approach to Contour Detection. ICCR, pp. 183–193.
Julesz, B., 1962. Visual pattern discrimination. IRE Trans. Inf. Theory 8, 84–92.
Kapadia, M.K., Ito, M., Gilbert, C.D., Westheimer, G., 1995. Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. Neuron 15, 843–856.

Kapadia, M.K., Westheimer, G., Gilbert, C.D., 1999. Dynamics of spatial summation in primary visual cortex of alert monkeys. Proc. Natl Acad. Sci. USA 96, 12073–12078.
Kapadia, M.K., Westheimer, G., Gilbert, C.D., 2000. Spatial distribution of contextual interactions in primary visual cortex and in visual perception. J. Neurophysiol. 84, 2048–2062.
Kastner, S., Nothdurft, H.C., Pigarev, I.N., 1999. Neuronal responses to orientation and motion contrast in cat striate cortex. Vis. Neurosci. 16, 587–600.
Knierim, J.J., van Essen, D.C., 1992. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J. Neurophysiol. 67 (5), 961–980.
Kovacs, I., Julesz, B., 1993. A closed curve is much more than an incomplete one: effect of closure in figure–ground segmentation. Proc. Natl Acad. Sci. USA 90, 7495–7497.
Kruizinga, P., Petkov, N., 1999. Nonlinear operator for oriented texture. IEEE Trans. Image Process. 8 (10), 1395–1407.
Kveraga, K., Ghuman, A.S., Bar, M., 2007. Top–down predictions in the cognitive brain. Brain Cogn. 65, 145–168.
La Cara, G.E., Ursino, M., 2008. A model of contour extraction including multiple scales, flexible inhibition and attention. Neural Netw. 21, 759–773.
Landy, M.S., Graham, N., 2004. Visual Perception of Texture. In: Chalupa, L.M., Werner, J. S. (Eds.), The Visual Neurosciences. MIT Press, Cambridge, MA, pp. 1106–1118.
Lamme, V.A.F., 1995. The neurophysiology of figure–ground segregation in primary visual cortex. J. Neurosci. 15, 1605–1615.
Lamme, V.A.F., Super, H., Spekreijse, H., 1998. Feedforward, horizontal, and feedback processing in the visual cortex. Curr. Opin. Neurobiol. 8, 529–535.
Lehky, S.R., Sereno, A.B., 2007. Comparison of shape encoding in primate dorsal and ventral visual pathways. J. Neurophysiol. 97, 307–319.
Li, C.Y., 1996. Integration field beyond the classical receptive field: organization and functional properties. News Physiol. Sci. 11, 181–186.
Li, C.Y., Lei, J.J., Yao, H.S., 1999. Shift in speed selectivity of visual cortical neurons: a neural basis of perceived motion contrast. Proc. Natl Acad. Sci. USA 96, 4052–4056.
Li, C.Y., 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.
Li, W., Li, C.Y., 1995. Shape and extent of the integration field of cat's striate cortical neurons. Acta Physiol. Sin. 47 (2), 111–119.
Li, W., Piech, V., Gilbert, C.D., 2006. Contour saliency in primary visual cortex. Neuron 50, 951–962.
Li, Z.P., 1998. A neural model of contour integration in the primary visual cortex. Neural Comput. 10, 903–940.
Li, Z.P., 1999. Contextual influences in V1 as a basis for pop out and asymmetry in visual search. Proc. Natl Acad. Sci. USA 96, 10530–10535.
Li, Z.P., 2002. A saliency map in primary visual cortex. Trends Cogn. Sci. 6 (1), 9–16.
Liang, K.H., Tjahjadi, T., Yang, Y.H., 1999. Bounded diffusion for multiscale edge detection using regularized cubic B-spline fitting. IEEE Trans. Syst. Man Cybern. 29 (2), 291–297.
Lindeberg, T., 1998. Edge detection and ridge detection with automatic scale selection. Int. J. Comp. Vis. 30, 117–156.
Likova, L.T., Tyler, C.W., 2008. Occipital network for figure/ground organization. Exp. Brain Res. 189, 257–267.
Long, L., Li, Y.J., 2008. Contour detection based on the property of orientation selective inhibition of non-classical receptive field. IEEE CIS 2008, 1002–1006.
Maffei, L., Fiorentini, A., 1976. The unresponsive regions of visual cortical receptive fields. Vis. Res. 16, 1131–1139.
Mallat, S.G., 1989. Multifrequency channel decompositions of images and wavelet models. IEEE Trans. Acoust. Speech Signal Process. 37 (12), 2091–2110.
Marr, D., Poggio, T., 1979. A computational theory of human stereo vision. Proc. R. Soc. Lond. B Biol. Sci. 204 (1156), 301–328.
Marreiros, A.C., Daunizeau, J., Kiebel, S.J., Friston, K.J., 2008. Population dynamics: variance and the sigmoid activation function. Neuroimage 42 (1), 147–157.
Maunsell, J.H., Nealey, T.A., DePriest, D.D., 1990. Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. J. Neurosci. 10, 3323–3334.
McGuire, B.A., Gilbert, C.D., Rivlin, P.K., Wiesel, T.N., 1991. Targets of horizontal connections in macaque primary visual cortex. J. Comp. Neurol. 305, 370–392.
McLean, J., Palmer, L.A., 1989. Contribution of linear spatiotemporal receptive field structure to velocity selectivity of simple cells in area 17 of cat. Vis. Res. 29, 675–679.
Mehrotra, R., Namuduri, K.R., Ranganathan, N., 1992. Gabor filter-based edge detection. Pattern Recognit. 25 (12), 1479–1494.
Menz, M.D., Freeman, R.D., 2003. Stereoscopic depth processing in the visual cortex: a coarse-to-fine mechanism. Nat. Neurosci. 6 (1), 59–65.
Mermillod, M., Guyader, N., Chauvin, A., 2005. The coarse-to-fine hypothesis revisited: evidence from neuro-computational modeling. Brain Cogn. 57, 151–157.
Moran, R.J., Kiebel, S.J., Stephan, K.E., Reilly, R.B., Daunizeau, J., Friston, K.J., 2007. A neural-mass model of spectral responses in electrophysiology. Neuroimage 37, 706–720.
Morrone, M.C., Burr, D.C., 1997. Capture and transparency in coarse quantized images. Vis. Res. 37 (18), 2609–2629.
Movshon, J.A., Newsome, W.T., 1996. Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. J. Neurosci. 16, 7733–7741.
Nelson, J.I., Frost, B.J., 1985. Intracortical facilitation among co-oriented, co-axially aligned simple cells in cat striate cortex. Exp. Brain Res. 61, 54–61.
Nothdurft, H.C., Gallant, J.L., van Essen, D.C., 1999. Response modulation by texture surround in primate area V1: correlates of “popout” under anesthesia. Vis. Neurosci. 16, 15–34.
Ohki, K., Reid, R.C., 2007. Specificity and randomness in the visual cortex. Curr. Opin. Neurobiol. 17, 401–407.
Oliva, A., Schyns, P.G., 1997. Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. Cogn. Psychol. 34, 72–107.
Ozeki, H., Finn, I.M., Schaffer, E.S., Miller, K.D., Ferster, D., 2009. Inhibitory stabilization of the cortical network underlies visual surround suppression. Neuron 62, 578–592.
Ozeki, H., Sadakane, O., Akasaki, T., Naito, T., Shimegi, S., Sato, H., 2004. Relationship between excitation and inhibition underlying size tuning and contextual response modulation in the cat primary visual cortex. J. Neurosci. 24, 1428–1438.
Papari, G., Campisi, P., Petkov, N., Neri, A., 2007. A biologically motivated multi-resolution approach to contour detection. EURASIP J. Appl. Signal Process. 1, 119-119.
Papari, G., Petkov, N., 2008. Adaptive pseudo-dilation for Gestalt edge grouping and contour detection. IEEE Trans. Image Process. 17 (10), 1950–1962.
Parker, D.M., Lishman, J.R., Hughes, J., 1997. Evidence for the view that temporospatial integration in vision is temporally anisotropic. Perception 26 (9), 1169–1180.
Peterhans, E., von der Heydt, R., 1989. Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. J. Neurosci. 9, 1749–1763.
Petkov, N., Westenberg, M.A., 2003. Suppression of contour perception by band-limited noise and its relation to non-classical receptive field inhibition. Biol. Cybern. 88, 236–246.
Pettet, M.W., McKee, S.P., Grzywacz, N.M., 1998. Constraints on long-range interactions mediating contour integration. Vis. Res. 38, 865–879.
Polat, U., Mizobe, K., Pettet, M.W., Kasamatsu, T., Norcia, A.M., 1998. Collinear stimuli regulate visual responses depending on cell's contrast threshold. Nature 391, 580–584.
Polat, U., Sagi, D., 1994. The architecture of perceptual spatial interactions. Vis. Res. 34, 73–78.
Ren, X.F., 2008. Multi-Scale Improves Boundary Detection in Natural Images. ECCV 2008. Springer, pp. 533–545.
Rockland, K.S., Lund, J.S., 1983. Intrinsic laminar lattice connections in primate visual cortex. J. Comp. Neurol. 216, 303–318.
Rodrigues, J., du Buf, J.M., 2006. Multi-scale keypoints in V1 and beyond: object segregation, scale selection, saliency maps and face detection. BioSystems 86, 75–90.
Ross, W.D., Grossberg, S., Mingolla, E., 2000. Visual cortical mechanisms of perceptual grouping: interacting layers, networks, columns, and maps. Neural Netw. 13, 571–588.
Rossi, A.F., Desimone, R., Ungerleider, L.G., 2001. Contextual modulation in primary visual cortex of macaques. J. Neurosci. 21 (5), 1698–1709.
Sceniak, M.P., Hawken, M.J., Shapley, R., 2001. Visual spatial characterization of macaque V1 neurons. J. Neurophysiol. 85, 1873–1887.
Sceniak, M.P., Ringach, D.L., Hawken, M.J., Shapley, R., 1999. Contrast's effect on spatial summation by macaque V1 neurons. Nat. Neurosci. 2, 733–739.
Series, P., Lorenceau, J., Fregnac, Y., 2003. The “silent” surround of V1 receptive fields: theory and experiments. J. Physiol. Paris 97, 453–474.
Shapley, R., Hawken, M., Ringach, D.L., 2003. Dynamics of orientation selectivity in the primary visual cortex and the importance of cortical inhibition. Neuron 38, 689–699.
Shen, Z.M., Xu, W.F., Li, C.Y., 2007. Cue-invariant detection of centre–surround discontinuity by V1 neurons in awake macaque monkey. J. Physiol. 583 (2), 581–592.
Shouval, H.Z., Goldberg, D.H., Jones, J.P., Beckerman, M., Cooper, L.N., 2000. Structured long-range connections can provide a scaffold for orientation maps. J. Neurosci. 20, 1119–1128.
Sigman, M., Cecchi, G.A., Gilbert, C.D., Magnasco, M.O., 2001. On a common circle: natural scenes and gestalt rules. Proc. Natl Acad. Sci. USA 98, 1935–1940.
Skottun, B.C., 1998. A model for end-stopping in the visual cortex. Vis. Res. 38, 2023–2035.
Sillito, A.M., Grieve, K.L., Jones, H.E., Cudeiro, J., Davis, J., 1995. Visual cortical mechanisms detecting focal orientation discontinuities. Nature 378, 492–496.
Somers, D.C., Nelson, S.B., Sur, M., 1995. An emergent model of orientation selectivity in cat visual cortical simple cells. J. Neurosci. 15, 5448–5465.
Somers, D.C., Todorov, E.V., Siapas, A.G., Toth, L.J., Kim, D.S., Sur, M., 1998. A local circuit approach to understanding integration of long-range inputs in primary visual cortex. Cereb. Cortex 8, 204–217.
Stepanyants, A., Martinez, L.M., Ferecskó, A.S., Kisvárday, Z.F., 2009. The fractions of short- and long-range connections in the visual cortex. Proc. Natl Acad. Sci. USA 106, 3555–3560.
Tang, Q.L., Sang, N., Zhang, T.X., 2007. Extraction of salient contours from cluttered scenes. Pattern Recognit. 40 (11), 3100–3109.
Ts'o, D.Y., Gilbert, C.D., Wiesel, T.N., 1986. Relationships between horizontal interactions and functional architecture in the cat striate cortex as revealed by cross-correlation analysis. J. Neurosci. 6 (4), 1160–1170.
Ursino, M., La Cara, G.E., 2004. A model of contextual interactions and contour detection in primary visual cortex. Neural Netw. 17, 719–735.
Walker, G.A., Ohzawa, L., Freeman, R.D., 2002. Disinhibition outside receptive fields in the visual cortex. J. Neurosci. 22, 5659–5668.
Walker, G.A., Ohzawa, I., Freeman, R.D., 1999. Asymmetric suppression outside the classical receptive field of the visual cortex. J. Neurosci. 19 (23), 10536–10553.
Walker, G.A., Ohzawa, I., Freeman, R.D., 2000. Suppression outside the classical cortical receptive field. Vis. Neurosci. 17, 369–379.
Weliky, M., Kandler, K., Fitzpatrick, D., Katz, L.C., 1995. Patterns of excitation and inhibition evoked by horizontal connections in visual cortex share a common relationship to oriented columns. Neuron 15, 541–552.

Wertheimer, M., 1923. Untersuchungen zur Lehre von der Gestalt. Psychol. Forsch. 4, 301–350.
Xu, W.F., Shen, Z.M., Li, C.Y., 2005. Spatial phase sensitivity of V1 neurons in alert monkey. Cereb. Cortex 15 (11), 1697–1702.
Yao, H.S., Li, C.Y., 1998. Symmetry and spatial summation properties of the integration field of cat's striate cortical neurons. Acta Biophys. Sin. 14 (3), 493–500.
Yazdanbakhsh, A., Livingstone, M.S., 2006. End stopping in V1 is sensitive to contrast. Nat. Neurosci. 9, 697–702.
Yen, S., Finkel, L.H., 1998. Extraction of perceptually salient contours by striate cortical networks. Vis. Res. 38, 719–741.
Zhaoping, L., Dayan, P., 2006. Pre-attentive visual selection. Neural Netw. 19 (9), 1437–1439.