A bio-inspired contour detection model using multiple cues inhibition in primary visual cortex

Chuan Lin[1] · Ze-Qi Wen[1] · Gui-Li Xu[2] · Yi-Jun Cao[1] · Yong-Cai Pan[1]

© Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

The human visual system has efficient architecture for information reception and integration for effectively performing visual tasks like detecting contours. Physiological evidence has shown that most neuronal responses in the classical receptive field (CRF) of the primary visual cortex are modulated, generally suppressed by the non-CRF surround. These center-surround interactions are thought to inhibit or facilitate responses to edges according to other similar edges in the surroundings, which is useful for suppressing textures and enhancing contours. A biologically motivated model with subfield-based inhibition is proposed in this paper to improve the performance of perceptually salient contour detection relative to the existing single-neuron based inhibition model. A novel subfield based inhibition framework is presented, where the inhibition terms are combined with center-surround and surround-surround differences using multiple cues, including orientation based energy distribution and directional saliency within regions. Extensive experimental evaluation demonstrates that the proposed method outperforms most of competing methods, especially biological motivated ones.

1 Introduction

Contour detection plays a fundamental role in the field of computer vision applications such as image segmentation and recognition [40] . In these applications, a contour normally represents a change in pixel ownership from one object or surface to another, which can be differentiated

✉ Chuan Lin
chuanlin@gxust.edu.cn

1   College of Electric and Information Engineering, Guangxi University of Science and Technology, Liuzhou 545006, China

2   Nanjing University of Aeronautics and Astronautics, Yudao Street 29, Nanjing, China

from edges that are most often defined as abrupt changes in low-level image features. One method of increasing the performance of contour detection involves integrating the local texture information into the context features. However, many researchers are involved in extensive efforts to combine a wider range of visual cues such as brightness, color, and texture information in pursuit of a contour detection model that provides human-level performance.

The study shows that human visual system evolved so as to be able to extract contour feature rapidly and effectively, such as, Landy et al. [32] studied the visual perception of texture and proposed Current Models of Texture Segregation. In order to address the question of how the brain integrates individual elements to construct the visual experience. Bar et al. [6] proposed a testable model for the rapid use of contextual associations in recognition. Albright et al. [1] summarized some information about use of contextual information to recover visual scene properties lost owing to the superimposition of causes in the visual image.

1.1 Related work

1.1.1 Non-biologically based models

Early contour detection approaches focused on finding local discontinuities, normally brightness, in image features. The Prewitt [43] and Sobel [21] operators detect edges by convolving a grayscale image with local derivative filters. Marr and Hildreth [37] used zero crossings of the Laplacian of Gaussian operator to detect edges. The Canny detector [10] also computes the gradient magnitude in the brightness channel, adding post-processing steps including non-maximum suppression and hysteresis thresholding. These methods can be improved by considering the combination of quadrature pairs of even and odd symmetric filters stabilized in a selected scale and orientation, such as the Gabor, log-Gabor and Gaussian derivatives [8] .

More recent approaches account for multiple image features, such as color and texture information, and use statistical and learning techniques [3, 39], active contours [30] and graph theory [27] (for a review of additional methods, see [40]). Martin et al. [39] built a statistical framework for difference of brightness (BG), color (CG), and texture (TG) channels, and used these local cues as inputs for a logistic regression classifier to predict the probability of boundary (Pb). Dollar et al. [20] proposed a Boosted Edge Learning (BEL) algorithm which, unlike the Pb method's reliance on such hand-crafted cues, aimed to learn and create a probabilistic boosting tree classifier to detect contours by using thousands of simple features computed on image patches. In order to make full use of global visual information, Arbelaez et al. [3] proposed a global Pb (gPb) algorithm designed to extract contours from global information by using the eigenvectors obtained from spectral partitioning. Salient contours were also extracted by building the Oriented Watershed Transform general post-processing method and Ultrametric Contour Maps.

In recent years, convolutional neural networks (CNNs) have become widely used in the field of computer vision and machine learning for applications such as classification [31], object detection [23], and contour detection [7, 36, 46] . Instead of the low-level image cues extracted in previous methods, CNNs can efficiently extract high-level features to improve contour detection performance. Shen et al. [46] customized a CNNs training strategy by partitioning contour (positive) data into subclasses and delineating each subclass using different model parameters. Bertasius el al. [7] exploited CNN based, object related features as high-level cues for contour detection, and built a Top-Down framework for detecting salient

boundaries. Xie and Tu [53] proposed a novel holistically nested architecture and used a deep supervision method to train the network.

1.1.2 Biological inspired model

Following another line of inquiry, there is a long history behind employing early visual mechanisms for image analysis. Some researchers focused on local statistics patterns [14] involving contrast dependence, orientation tuning and spatial asymmetry, and additionally employed other concepts used in visual applications, such as contour [55, 57], color constancy [58] and saliency [59] detection.

In recent decades, many biologically motivated contour detection models have been proposed and showed good performance for gray-scale natural images [5, 24, 35, 41, 47–49, 51, 56, 60, 61] . Grigorescu et al. [24] exploited a novel framework of center-surround interactions, in which simple and complex cells were stimulated by Gabor filters to mimic the center region response. Furthermore, a linear surround inhibition approach was used to model the responses of the surround region, creating texture suppression with a distance cue. Following this framework, several improved models have been proposed. Some of them modified the inhibition region architecture by developing a butterfly-shape rather than ring-shape inhibition region [41, 48, 60, 61], dividing the inhibition region into four parts: two side and two end regions along an optimal orientation using both end regions for facilitation [60] and either unilateral [61] or both side regions [41, 48, 60] for suppression. Others have aimed to exploit valuable and efficient cues in combination, including orientation [35, 48], energy distribution and luminance contrast [56], multi-scales [51], luminance [49, 56], sparseness [47], and color [55, 57] .

1.2 Center-surround interactions in the primary visual cortex (V1)

Orientation significance based edge perception is one of the fundamental functions of the human visual system (HVS). The groundbreaking work of Hubel and Wiesel [25] reveals that the majority of neurons in the V1 do not respond to diffuse light, whereas they are extremely sensitive to oriented bars placed in a limited area of view, named the classical receptive field (CRF). Later, extensive neurophysiological evidence [13, 50] showed that a surrounding region beyond the CRF, called the non-CRF or surround, is able to modulate and in most cases suppress the reaction of the CRF to stimulation, although unrelated stimulus of the non-CRF region does not activate these neurons.

The non-CRF modulation, called center-surround interaction, allows neurons to integrate information from relatively large parts of the visual field. Many neurophysiological findings indicate that the degree of non-CRF inhibition depends on contrasts in the center-surround texture pattern, such as orientation [18], luminance [45], spatial frequency [9], spatial phase [54] and relative moving speed [34] . From a computational modeling viewpoint, the center-surround mechanism integrates spatial contextual information and helps to improve the performance of many visual tasks, such as contour detection [5, 24, 35, 41, 47–49, 51, 56, 60, 61], visual attention [28], image segmentation [52], and color constancy [22] .

Figure 1a shows the structure of a previous bio-motivated contour detection model. Models such as this one compute center region responses and make use of the suppressed modulation of neurons (pixels) in the surround region. However, most neurons in the V1 exhibit different functions and spatial structures, especially the non-CRF region. Li et al. [33] found that the

different surround structures may be adapted to analyze different texture patterns in visual images. Moreover, signals from the surround have been reported both to suppress and facilitate responses, selectively and unselectively [12] . Additional studies on the non-CRF [12, 50] found that suppression may originate from any specific localized region of the surround. This implies the inhibitory interactions can take different forms, such as a subfield distributed in the surround region. As shown in Fig. 1b, the center region is modulated by a subfield distributed in the surround.

As for the possible biological mechanisms underlying the neural 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 [44] . In addition, physiological and psychophysical results indicate that feedback projections from higher cortical areas (e.g., V2, V4) and subcortical mechanisms [2], also provide clear substrate for neural interactions in V1. However, in this paper, we focus only on the functionalities of neural interactions rather than on the detailed connection method in neuroscience.

1.3 Contributions of the proposed model

Most previous bio-motivated contour detection models use a family of Gabor filters [19, 29] to obtain the preliminary responses of contours, then subtract pixels based on surround inhibition terms with a distance weight simulated using a Difference of Gaussian (DoG) function. This work aims to utilize novel subfields, rather than pixels, in the surround region. The remainder of this section briefly describes the main contributions of this paper.

First, a subfield-based inhibition framework is proposed. These subfields can receive more information from neighboring areas than previous pixel-based methods. A linear combination of subfields was used to compute the new inhibition term.

Second, subfield inhibition cues were utilized, including energy distribution in different orientations and directional saliency. Two approaches were used to compute inhibition responses: (1) distribution differences between the center and subfield regions multiplied by the central directional saliency, named the center-surround interaction; (2) distribution differences between multiple subfields, named the surround-surround interaction.
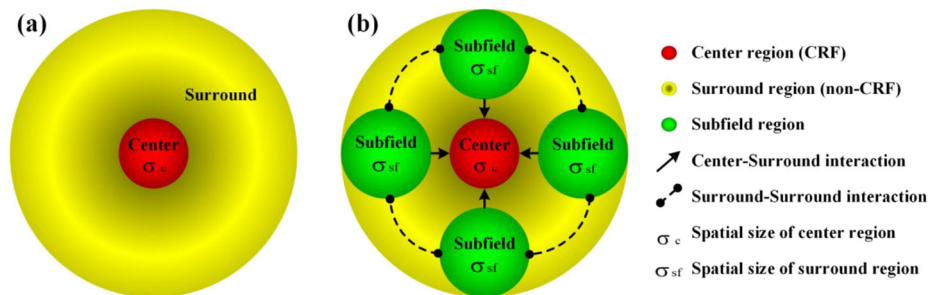

Fig. 1 The non-CRF structure of the previous and proposed subfield-based model. **a** Traditional pixels-based surround suppression model. Each neurons (pixels) in surround region (yellow region) will inhibit center neurons (pixels). **b** Proposed subfield-based surround suppression model. Subfield (green regions) instead of individual pixels to take part in suppression computation. Dotted and solid line indicates surround-surround and center-surround interaction, respectively

The remainder of this paper is organized as follows: Section II describes the proposed contour detection model in detail, Section III presents a performance evaluation of the proposed model using RuG and BSDS300/500 datasets, and Section IV contains the discussion and conclusion of this work.

2 Contour detection

2.1 Overview

The HVS is an efficient code/decoder machine for visual information processing. From retina photoreceptors to the V1, early visual neurons aim to extract salient local structure information such as extreme points and edges. The neurons in the retina perceive local information from external scenes by transforming optical signals into neural responses, then transmit this information to different neurons though feed forward or feedback connections. Most connections can be calculated as a set of neural filters to extract specific local features in limited spatial regions, such as the DoG function for lateral geniculate nucleus cells [17] and Gabor function for V1 neurons [19, 29] .
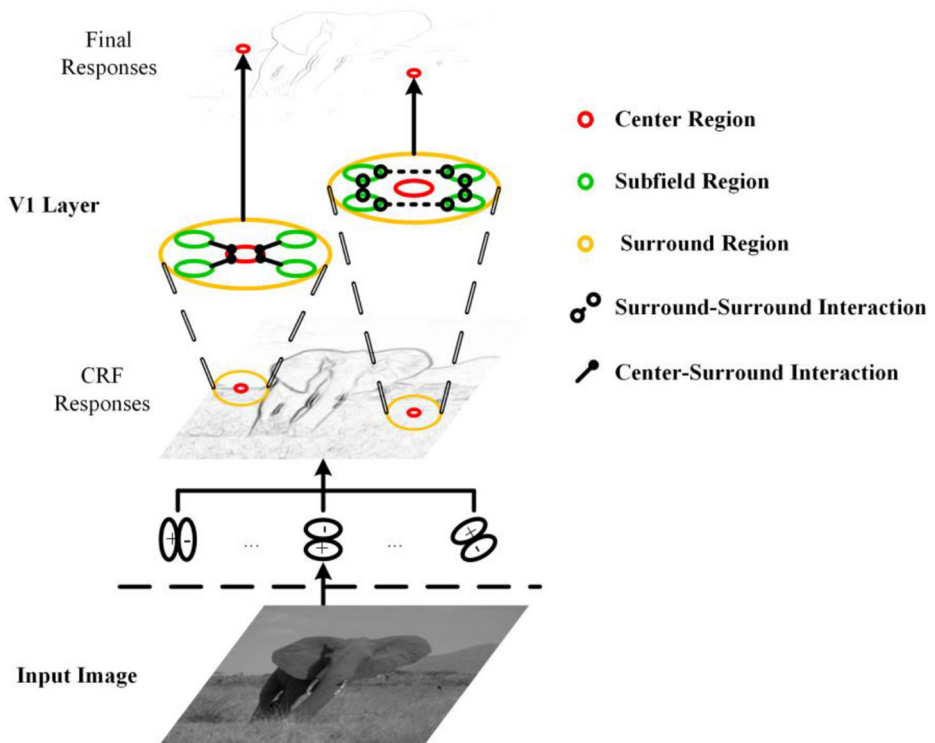
Fig. 2 Networking structures of proposed models. Each point in the CRF responses map is suppressed by its subfield region (green circle). The left or right receptive field in the CRF responses map indicates center-surround inhibition method and surround-surround inhibition method, respectively

In this work, a new framework is proposed based on combined subfield inhibition with center-surround and surround-surround interactions for the specific task of contour detection. The general networking structures of the proposed model are shown in Fig. 2. The gradient magnitude is first computed along multiple orientations; these are then combined as the responses of the CRF in the V1 region. Subsequently, the surrounding subfield in the non-CRF region is built to participate in the computation of non-CRF responses. In this step, two inhibition cues are extracted from the responses of CRF (energy distribution and directional saliency), and then the corresponding inhibition weights of each subfield are computed based upon the differences of center-surround and surround-surround interactions. In the concluding step, non-CRF responses $Inh(x, y)$ are subtracted from CRF responses $CRF(x, y; \sigma)$ at each location to obtain the final responses:

$$R(x,y) = H(CRF(x,y;\sigma) - \alpha Inh(x,y)), \tag{1}$$

where $H(x) = max (0, x)$ is used to guarantee that neuronal responses will not be negative. The factor α denotes the inhibition strength between the neurons within the CRF and its surrounding non-CRF, and we set $\alpha = 0.5$.

The detailed CRF responses will be descripted in Section II.B. Section II.C to Section II.E will, respectively, describe the division of subfields, the subfields features and how to compute the final inhibition terms with center- and surround-surround methods.

2.2 Responses of CRF in V1

The derivative of the 2D Gaussian function may be used to describe the spatial summation properties of neurons in the CRF. The derivative of the 2D Gaussian function can be expressed as follows:

$$GD(x,y;\theta,\sigma) = \frac{-\hat{x}}{\sigma^3\sqrt{2\pi}}\exp\left(-\frac{\hat{x}^2 + \gamma^2\hat{y}^2}{2\sigma^2}\right), \tag{2}$$

where $\hat{x} = xcos(\theta) + ysin(\theta)$, $\hat{y} = -xsin(\theta) + ycos(\theta)$, θ is the preferred orientation of a neuron, the spatial aspect ratio γ determines the elliptical nature of the CRF, and σ presents the size of the derivative of the Gaussian kernel. In this paper, $\gamma = 0.5$ on the basis of physiological findings [29] .

For an input image $I(x, y)$, the neural responses of the CRF with preferred orientation θ_i are defined by a convolutional operator:

$$e_i(x,y;\theta_i,\sigma) = |I(x,y) \star GD(x,y;\theta,\sigma)|, \tag{3}$$

At each location, the neural responses e_i can be calculated using a different orientation θ_i with a number N_θ:

$$\theta_i = \frac{(i-1)\pi}{N_\theta}, i = 1, 2, , N_\theta, \tag{4}$$

in this model, $N_\theta = 8$.

Then, at each location, a winner-take-all strategy is performed such that the maximum CRF response over all N_θ along different orientations is selected as the final CRF response, written as

$$CRF(x, y; \sigma) = \max\{e_i(x, y; \theta_i, \sigma)|, i = 1, 2, , N_\theta\}.$$ (5)

2.3 Subfield properties

The proposed subfield based inhibition model was inspired by multiple physiological findings [13, 50] on neural connections creating intracortical inhibition. The suppressing connections between neurons in V1 are composed of many individual subfields, some of these which transmit effects to other regions. The main difficulties for modelling distribution and subfield properties are listed below:

- Spatial distribution of subfields in the non-CRF region.
- Complexity of proposed computational model for contour detection.
- Local cues in each subfield and combining approaches between them.

The non-CRF region was built as a series of feature-receptive units, named subfields, to obtain local low-level cues. In this model, each included subfield is represented in parametric form by a tuple (r, d, ϕ) where the parameter r represents the scale (radius) of a subfield while d and ϕ indicate the distance and the bias angle between the CRF and subfield centers, respectively. The set of 3-tuples that represent the configured subfield above are denoted as $S_i = \{(r_s^i, d_s^i, \phi_s^i)|i = 1, , N_S\}$.

The spatial distribution of the subfields must maintain flexibility and integrity for extracting local features, as well as lowering the complexity as much as possible. To balance the complexity and performance, subfields were set only within a circle around
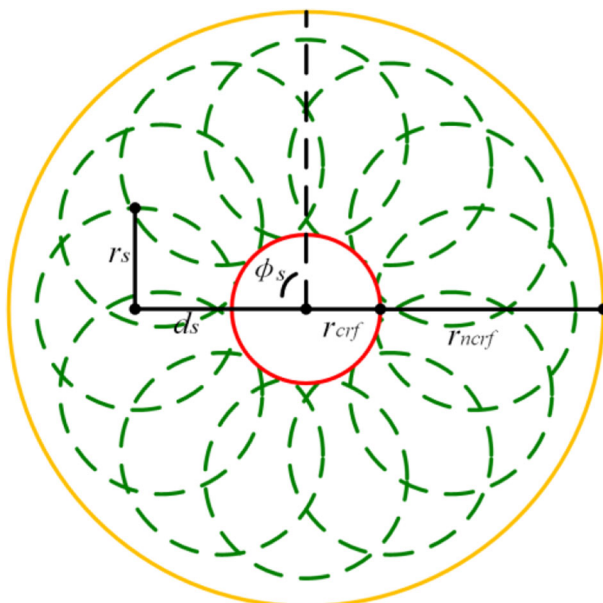

Fig. 3 The specific architecture of subfield in non-CRF

CRF region, and N_S was experimentally set to 12. In this model, the pixel radius of the CRF and non-CRF regions were set to $r_{crf} = 2.5\sigma$ and $r_{ncrf} = 7.5\sigma$, respectively. The parameters of a set S were specified as $r_s^i = 2\sigma$, $d_s^i = 4\sigma$, and $\phi_s^i = \frac{2\pi(i-1)}{N_S}$ (the counterclockwise angle starting from the positive y-axis axis) for $i = 1, \ldots, N_S$. As shown in Fig. 3, the subfields approximately cover the entire non-CRF region without excessive overlap.

2.4 Inhibition feature of subfield and center region

In this subsection we will focus on constructing a feature cue that can represent the characteristics of the subfield and center region.

Orientation selectivity is one of the most fundamental properties in V1 cells. Some neurophysiology studies [26] discovered that many cortical cells are spatially arranged within the visual cortex according to similar orientation characteristics, and this functional architecture is refer to as an orientation column. At the cellular level, all possible directions in the certain CRF can be characterized and fused in an orientation column. In computation, we can see the orientation column as a vector which represents energy distribution with all of the orientations.

The energy distribution $\overrightarrow{E}(x, y)$ is defined as [56]:

$$\overrightarrow{E}(x, y) = [\overline{e}_1, \overline{e}_2, , \overline{e}_{N_\theta}](x, y), \tag{6}$$

where $\overline{e}_i$ denotes the CRF responses e_i after Gaussian blurring with scale σ.

This vector involves local energy of orientations N_θ from a point (x, y) which will be used to construct regional distributions and compute the difference in stimuli between the CRF and subfield. For an individual region, the energy distribution is computed as the pointwise sum in the region before normalization. Let $\overrightarrow{E}_c(x, y)$ and $\overrightarrow{E}_s(x, y)$ denote, respectively, the energy distributions of the CRF center and the $S_i th$ subfield. These are computed as

$$\overrightarrow{E}_c(x, y) = \sum\nolimits_{(x,y) \in A_{crf}} \overrightarrow{E}(x, y) / N_{A_{crf}}, \tag{7}$$

$$\overrightarrow{E}_{s_i}(x, y) = \sum\nolimits_{(x,y) \in A_{sf_i}} \overrightarrow{E}(x, y) / N_{A_{sf_i}}, \tag{8}$$

where A_{crf} and $N_{A_{crf}}$ represent the CRF (center) region and the number of points in A_{crf}, respectively; and A_{sf_i} and $N_{A_{sf_i}}$ represent the i-th subfield region and the number of points in A_{sf_i}, respectively. As shown in Fig. 3, A_{crf} is the region surrounded by red solid lines and A_{sf_i} is the region surrounded by green dotted lines.

2.5 Non-CRF responses using center-surround and surround-surround inhibition

In this subsection, the proposed models are described in detail by using cues from the local regions of each subfield, including adopted subfield to subfield (surround-surround) and subfield to center (center-surround) interactions.

2.5.1 Center-surround inhibition

The strength of surround inhibition decreases with increasing feature differences between the CRF and non-CRF [56] . To effectively represent the energy distribution between the texture patterns within CRF and subfield, these differences between the center and subfields are described as:

$$\Delta E(x,y) = \sum_{i=1}^{N_S} \left\| \vec{E}_c(x,y) - \vec{E}_{s_i}(x,y) \right\|_1, \tag{9}$$

where $\|\cdot\|_1$ denotes the L1-norm. Equation (9) describes the summation of differences between the center and all subfields of energy distribution in the non-CRF. Each difference can be defined as $\vec{E}_c(x,y)$ minus $\vec{E}_{s_i}(x,y)$, following a L1-norm operator.

Center-surround interactions in the proposed model are defined as the summation difference, $\Delta E(x, y)$, modulated by the directional saliency, $D_c(x, y)$:

$$CS(x,y) = \Delta E(x,y) \cdot D_c(x,y), \tag{10}$$

$$D_c(x,y) = \frac{1}{N_\theta} \sum_{i=1}^{N_\theta} \left(\vec{E}_c(x,y) - \mu_{E_c} \right)^2, \tag{11}$$

where μ_{E_c} is the average value of the distribution $\vec{E}_c(x,y)$. The term $\vec{E}_c(x,y) - \mu_{E_c}$ represent each element in $\vec{E}_c(x,y)$ minus μ_{E_c}.

$D_c(x, y)$ is an important visual feature for understanding natural scenes. Early physiology and neurophysiology studies [25, 26] have shown that most cortical cells in cats or monkeys do not respond to diffuse light, but have a strong response to specific azimuthal light stimulation. Subsequent research [12] revealed additionally evidence of orientation-selective cells existent in the CRF and non-CRF. It is reasonable that if the directionality of a region around a point is not significant, meaning bars in different directions exist in the region and cause lower directional saliency, the possibility of becoming a contour would be low. This is shown in Fig. 4b by the red and green solid line circles. Conversely, the probability of becoming a contour is high if higher directional saliency exists around a point, as shown in Fig. 4a by the red and green circles, and in Fig. 4b by the green dotted circle.

In Fig. 4, eight blue bars of a plot indicate energy distribution $\vec{E}(x,y)$ while the red bar is the degree of directional saliency $D_c(x, y)$. Both are clearly described in the properties of the local region, and prove helpful in textural suppression.

2.5.2 Surround-surround inhibition

In some cases, isolated edges are surrounded by cluttered backgrounds, and these edges tend to exhibit intensive responses. In such cases, the general center-surround inhibition term $CS(x, y)$ is insufficient for completely suppressing the edge response. Hence, the proposed model makes use of the surround-surround inhibition term $SS(x, y)$ for further texture suppression. The difference among all of subfields is used to compute $SS(x, y)$, from which the difference of multiple distributions can be decompose into two simple steps. First, the energy distribution of each subfield $\vec{E}_{s_i}(x,y)$ is combined into a variable using L1-norm; then, STD is used to
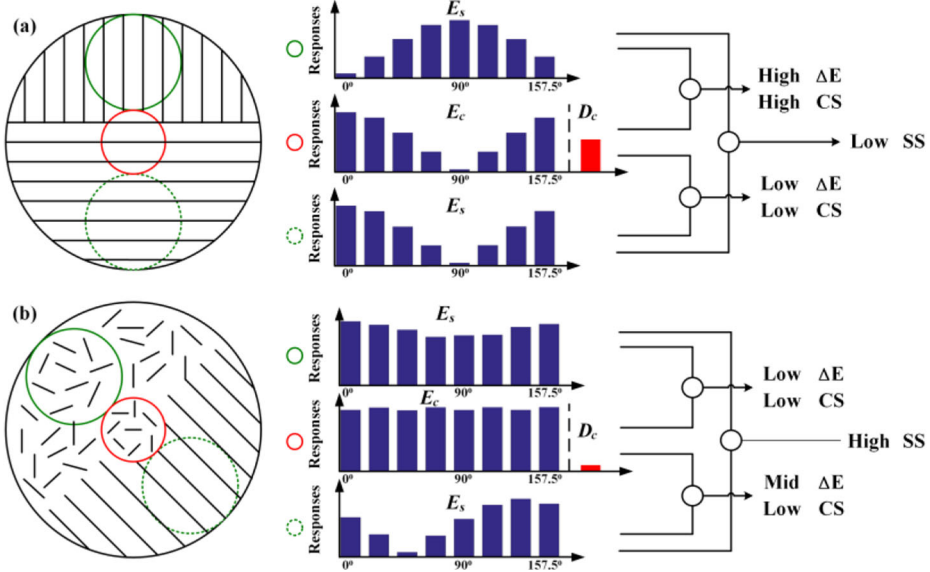
Fig. 4 Examples of computing the energy distribution and directional saliency for two typical texture patterns. In this paper, we define the horizontally leftward orientation as 0° and the vertically upward orientation as 90°. In the right panels of the figures, the blue bars represent the component of the energy distribution vector in center or subfield (i.e., $\vec{E}_c(x,y)$ or $\vec{E}_{s_i}(x,y)$); the red bars represent the directional saliency in center region (i.e., $D_c(x, y)$).

combine the variables. The process is performed as below:

$$SS(x,y) = \frac{1}{N_S}\sum_{i=1}^{N_S}\left(\left\|\vec{E}_{s_i}(x,y)\right\|_1 - \mu_{E_s}\right)^2, \tag{12}$$

where μ_{E_s} is the average of a set $\|E_{s_i}(x,y)\|_1$ for $i = 1, \ldots, N_S$. Higher $\|E_{s_i}(x,y)\|_1$ values mean the summation CRF responses in *ith* subfield are higher, and vice versa.

Examples of computing the center-surround and surround-surround inhibition on four and two types of texture patterns, respectively, are shown above in Fig. 4. ΔE is high when the stimuli within the CRF and subfield have different energy distributions, and low when the stimuli within the CRF and subfield have similar energy distributions. However, as seen in Fig. 4b, an existing disorder measurement CS that is lower than ΔE, implies larger inhibition strength. Furthermore, SS is high when the CRF responses among subfields have large differences, and SS is low when the CRF responses among subfields have a low difference pattern.

2.5.3 The final non-CRF responses

The final non-CRF response (inhibition term) is obtained by combining traditional distance weight and linear integration of $\overline{CS}$ and $\overline{SS}$. Their formation is shown as:

$$Inh(x,y) = W_d \cdot \left(\overline{CS}(x,y) + \overline{SS}(x,y)\right), \tag{13}$$

where $\overline{CS}(x,y)$ and $\overline{SS}(x,y)$, respectively, indicate the final inhibition terms of center-surround and surround-surround suppression, which are listed below:

$$\overline{CS}(x,y) = \exp\left(-\frac{\mathcal{N}(CS(\S,\dagger))}{2\sigma_{cs}{}^2}\right), \tag{14}$$

$$\overline{SS}(x,y) = \exp\left(-\frac{SS(x,y)}{2\sigma_{SS}{}^2}\right), \tag{15}$$

where $\mathcal{N}(\S) = \min(\infty, \max(\prime, \S))$ is a compound operator that ensures the output fits into [0, 1]. In this work, $\sigma_{cs} = 0.25$ and $\sigma_{SS} = 0.9$ were set to ensure the output value is as meaningful as possible.

The distance weight [24] is obtained by convolution operator with CRF responses $CRF(x, y; \sigma)$ and normalized different of Gaussian (DoG) function:

$$W_d = CRF(x,y;\sigma) \star \overline{DoG}(x,y;\sigma), \tag{16}$$

$$\overline{DoG} = \frac{\max(0, DoG(x,y;\sigma))}{\|\max(0, DoG(x,y;\sigma))\|_1}, \tag{17}$$

$$DoG(x,y;\sigma) = \frac{1}{2\pi(4\sigma)^2}e^{-\frac{x^2+y^2}{2(4\sigma)^2}} - \frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}, \tag{18}$$

3 Results

First in this section, the qualitative experiments of a natural image from a dataset (BSDS300) [38] and a synthetic image are given to intuitively exhibit the inhibition properties of the model. Next, results from model tests using the whole RuG40 [24], BSDS300 [38], and BSDS500 [3] datasets are shown alongside a comparison of the quantitative performance. Some parameters in our model are fixed ($\sigma_{cs} = 0.25$, $\sigma_{SS} = 0.9$), and the others are variable, such as scale of CRF kernel $\sigma \in \{2.0, 2.5, 3.0, 3.5\}$ and subfield configuration $r_s \in \{2, 3, 4, 5\}\sigma$, $d_s \in \{4, 6, 8\}\sigma$, and $N_s \in \{2, 4, 8, 12\}$.

3.1 Qualitative analysis

3.1.1 Basic properties of inhibitory effects

For a clear and detailed demonstration of the intuitive performance of the proposed model the inhibition terms and responses of the model were generated with individual center-surround and surround-surround inhibition effects as shown in Fig. 5. In this experiment, we set $\sigma = 2.0$, $r_s = 6\sigma$, $d_s = 6\sigma$, and $N_s = 12$. When computing the gradient magnitude map, it was found that many unwanted texture edges existed in the CRF response map (Fig. 5e). In order to obtain a better contour map without textures, the center-surround inhibition term $\overline{CS}$ (Fig. 5c), surround-surround inhibition term $\overline{SS}$
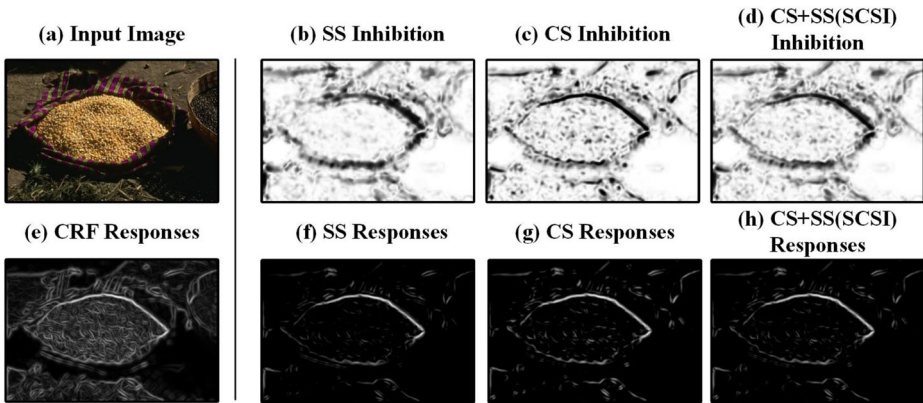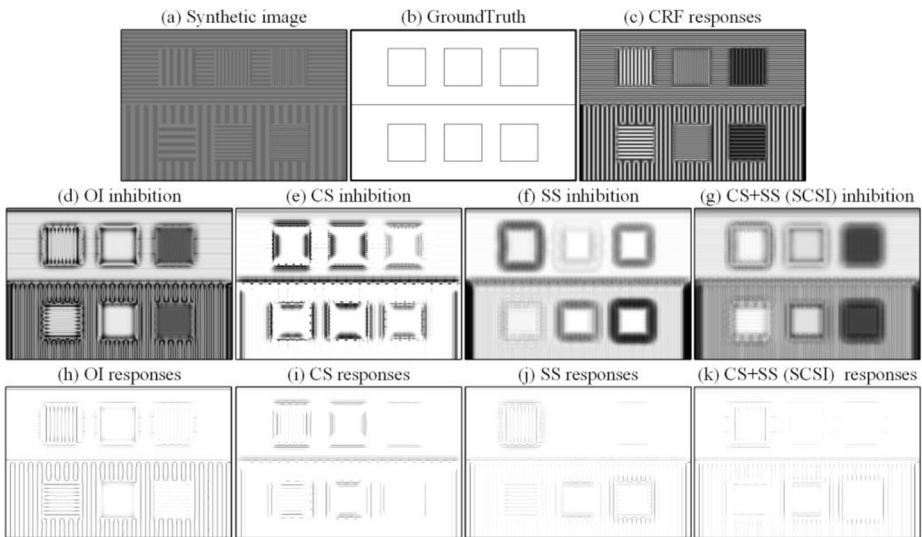
Fig. 5 Examples of inhibition term and contour responses maps. **a** input image, **b-d** the inhibition term modulated by surround-surround, center-surround and the combination methods, respectively, **e** the CRF responses, **f-h** the final responses to salient contours with such of the individual term Fig. 5b-d inhibition by replacing $Inh(x, y)$ in (1)

(Fig. 5b) and combination Inh (Fig. 5d) were used. The $\overline{CS}$ term (Fig. 5c) is seen to efficiently suppress texture, but poorly inhibits some dense textures (corn kernels in the image center) and salient useless edges embedded in the cluttered background. Conversely, the $\overline{SS}$ term (Fig. 5b) shows good inhibition in these respects. Consequently, the final inhibition term in Fig. Subfield- and pixel-based suppression models on synthetic image and $\overline{CS}$ for more efficient texture suppression and contour protection.

3.1.2 Subfield- and pixel-based suppression models on synthetic image

In this part, we will do a qualitative experiment of the difference between subfield- and pixel-based suppression models on a synthetic image, and intuitively explain where the subfield-model performs well and where it is not as good as the pixel-based model. In this experiment, we set $\sigma = 2.0$, $r_s = 6\sigma$, $d_s = 6\sigma$, and $N_s = 12$ for our method and the parameter combination in [56] for OI.

Figure 6a, b exhibit the synthetic image and corresponding hand-painted possible contour, respectively. The synthetic image consists of eight rectangular patterns with different widths and orientation stripes. Two of the patterns divide the image into two parts: (1) upper part with horizontal small width stripes; (2) lower part with vertical large width stripes. Six squares are embedded in the two-part image with stripes of gradually changing width. Figure 6c shows the CRF responses output from (5). Figure 6d, e exhibit center-surround inhibition responses with pixel-based models, named orientation inhibition (OI) [56], and subfield based, named center-surround inhibition (CS), which only uses the term $\overline{CS}(x, y)$ in Eq. (13) for final responses computation. Figure 6f exhibits the surround-surround inhibition response — using the term $\overline{SS}(x, y)$ instead of $\overline{CS}(x, y) + \overline{SS}(x, y)$ in Eq. (13), with the subfield-based model. Figure 6h–k show the final contour responses by subtracting corresponding inhibition terms (Fig. 6d–g) from CRF responses (Fig. 6c), respectively.

Fig. 6 Examples of inhibition and contour responses on a synthetic image. **a** synthetic image; **b** hand-painted possible contour; **c** CRF responses obtained from (5); **d**, **e** pixels based, subfield based center-surround inhibition responses, respectively; **f** subfield based surround-surround inhibition responses; **g** the inhibition term modulated by the combination methods; **h-k** the final contour responses by subtracting corresponding inhibition terms (**d-g**) from CRF responses (**c**), respectively

In order to fairly compare the advantages and disadvantages of pixel and subfield based model, we set the experiments under two conditions. First, we just used non-orientation selectively method and energy distribution feature $\vec{E}(x, y)$ to obtain inhibition term and the corresponding final response. Then, we used the same scale parameter $\sigma = 2$ and independent optimized suppression parameters.

Compared with the pixel-based OI model, the subfield-based CS model has an obvious advantage: robust textures suppression at different scales, such as the lower part with large scale textures and the leftmost two rectangles. However, the CS model is worse than the OI model in slight contour protection, such as the corner part of the rectangles. This is because the subfields can capture a wide range of region information in contrast to the pixel-based fine information.

From Fig. 6f, j, we can see that the SS model shows a good performance in contour protection at corners or in parts where the stripe widths differ greatly, such as the top left and bottom right rectangles. This is because (1) the SS model is only concerned with the contrast of the surrounding region; (2) the larger difference of stripe widths will result in a larger difference of energy density, and thus, according to (12) and (15), the inhibition magnitude of $\overline{SS}(x, y)$ will be decreased.

3.2 Quantitative experiment

3.2.1 RuG dataset

Model performance was measured using an available contour detection dataset, the RuG40 dataset [24], which includes 40 natural images and an associated ground truth drawn by hand.

This dataset has been widely used to evaluate the performance of biologically motivated contour detectors [5, 24, 35, 41, 47–49, 51, 56, 60, 61] . In order to quantitatively evaluate the performance of this method, binary contour maps were constructed using the standard procedure of non-maxima suppression followed by hysteresis thresholding. The similarity between the binary contour map detected by a model and the corresponding ground truth were measured, as defined in [24].

$$P = \frac{card(E)}{card(E) + card(E_{FP}) + card(E_{FN})}, \quad (19)$$

where $card(X)$ denotes the number of elements of the set X, E denotes the set of correctly detected contour pixels, and E_{FP} and E_{FN} are the sets of false positives (spurious contours) and false negatives (ground-truth contours missed by the contour detector), respectively. The performance measure P is a scalar able to take values in the interval [0, 1], where a higher measure P reflects better precision in contour detection. As performed in previous works [5, 24, 35, 41, 47–49, 51, 56, 60, 61], a contour pixel was considered to be correctly detected if a corresponding ground tru th contour pixel was present in a 5 × 5 square neighborhood centered at the respective pixel coordinates.

In this experiment, series of parameter combinations were selected for statistical analysis on the number N_s, radius r_s and distance d_s of subfield. We set $\sigma = 2.0$ and $r_s = [2, 3, 4, 5]\sigma = [4, 6, 8, 10]$, $d_s = [4, 6, 8]\sigma = [8,12,16]$, and $N_s = [2, 4, 8, 12]$ for combinations. In Fig. 7, the heights of the bars reflect the optimal dataset scale (ODS) of P over all 40 images in the RuG dataset with different parameter combinations. One parameter was fixed while the remaining two parameters were measured. The optimal performance P and the corresponding parameters over all images are listed at the top of each. As we see from the left two subplots, the model performance increases as the number of subfields increases, while from the right subplot the lower d_s makes r_s a monotonous decreasing whereas the larger d_s makes r_s a monotonous increasing. According to the results from the left two subplots, we choose $N_s = 12$. In addition, the result from the right subplot reveals that subfields should be close to the center CRF, if not, it should cover the non-CRF as far as possible. Based on this experiment, the subfield configuration in below measurements are used $r_s = 2\sigma$, $d_s = 4\sigma$, and $N_s = 12$.

We test our subfield based center-surround and surround-surround inhibition (SCSI) model at different noise levels to verify the capacity to extract contours in noised images. Figure 8 clearly shows that our model has ability of suppressing noises with increasing σ. When the
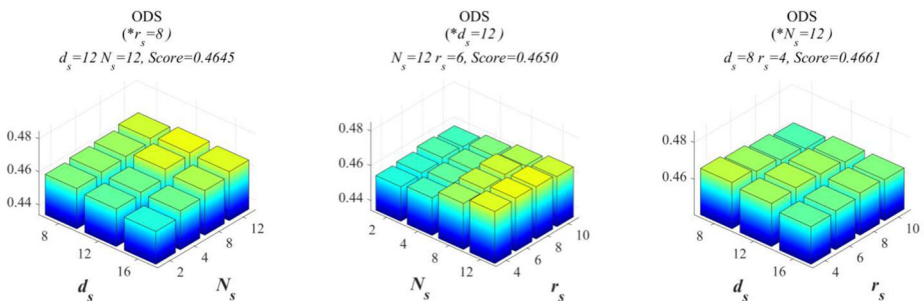
Fig. 7 Optimal dataset scale (ODS) of SCSI over all images of the RuG dataset with various parameter combinations and we fixed $\sigma = 2.0$. Each bar in the figure represents the ODS over all the images for a given parameter combination. The text above each panel highlights the best ODS, which is obtained using the given parameter values. Symbol '*' indicates the fixed parameters
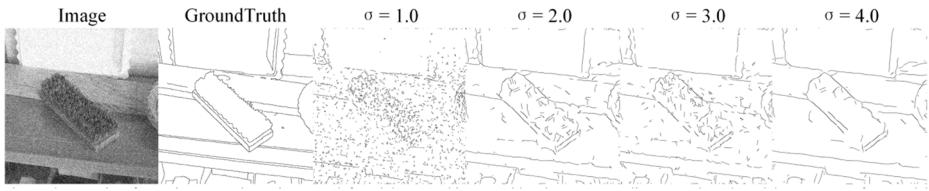
Fig. 8 The capacity of SCSI in suppressing noise. From left to right: natural images with noise, corresponding Ground Truth, and the responses of our model with different filter scales σ. Each response is constructed using the standard procedure of non-maxima suppression followed by hysteresis thresholding

scale σ is increased to proper values, the target contour can pop-out clearly from the noised background.

We evaluate and compare proposed SCSI with two other biologically inspired models, i.e., ISO-gd [56], OI [56] . Figure 9 shows three images, corresponding contour maps, and performances with measurement precision P. Note that the ISO-gd contour detection model mentioned in this paper presents the optimized version of the isotropic inhibition model [24] proposed in [56] . As shown, the surround inhibition mechanism in the *ISO-gd* and *OI* method can efficiently suppress the textures, though there remains room for improvement in its capacity for removing texture edges. In contrast, the *SCSI* is better than the others at texture suppression (e.g. stones on the river bank and dense groves around the bear and rhino), where it retains and enhances remarkable contours. The precision, P, clearly shows the performances of the three detectors on the best threshold.

In order to represent richer performance on the RuG database, multiple thresholds were used to generate a series of assistant measurement false positives (e_{FP}) and false negatives (e_{FN}) as defined in [24], and produce an $e_{FP} - e_{FN}$ curve. These were calculated as follows:

$$e_{FP} = card(E_{FP})/card(E), \tag{20}$$
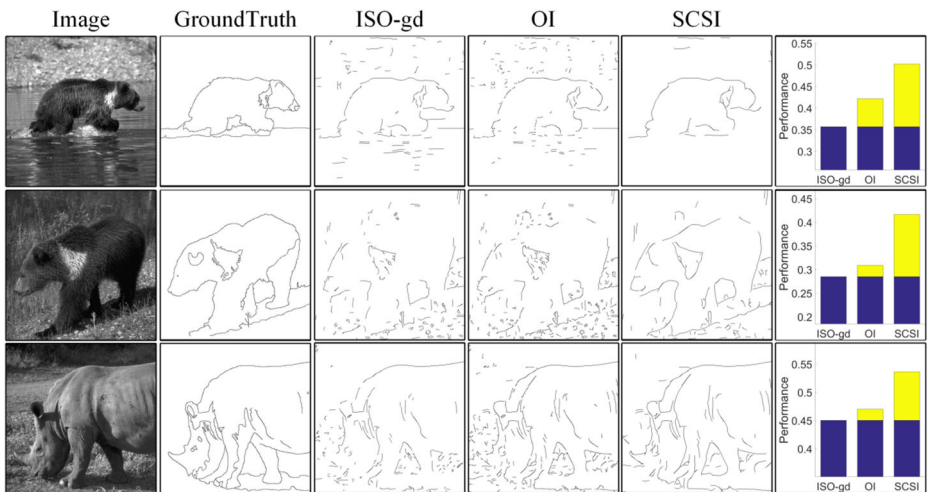
Fig. 9 Results of various models on three example images from the RuG40 dataset. The columns from left to right represent the original image, corresponding ground truth, ISO-gd, OI and SCSI detectors, the performance (P value) obtained from its right contour map (i.e., *ISO-gd*, *OI* and *SCSI*) using the optimal parameter setting, respectively

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

where E_{GT} denotes the set of contour pixels in the ground truth contour image, the indicator e_{FN} is a scalar taking values in the interval [0, 1], and e_{FP} is in the interval [0, +∞]. Lower values for both of these reflect better performance regarding false positives and false negatives, respectively. Each point on the curve is computed independently first by hysteresis thresholding the final contour map after non-maxima suppression to produce a binary boundary map, and then by matching this machine boundary map against each of the human boundary maps in the ground truth dataset.

Figure 10 shows the $e_{FP} - e_{FN}$ curve of the *SCSI* contour detection models. Note that lower e_{FP} and e_{FN} values indicate better performance of textures suppression and contour protection, respectively. In this figure, the lower e_{FP}, the better textures inhibition performance, while the lower e_{FN}, the better ability of contour protection. We can see that as the receptive field scale σ increases, the texture suppression ability of the model is stronger, whereas the protection ability of the contour is weaker. Additionally, it can be seen that *SCSI* show good performance with a scale parameter $\sigma = 2.0$. The proposed $e_{FP} - e_{FN}$ curve enriches the quantitative performance on the RuG data set. It can intuitively describe the degree of false positives and false negatives instead of relying on the single precision P proposed in [24] .

Table 1 shows a quantitative comparison of the P-measures of various models on the RuG images, including the *ISO-gd*, *Butterfly-shaped* [61], *MCI* [56], CSLCM [11] and proposed model ($\sigma = 2.0$). The optimal threshold is 0.06. Although the overall evaluation of our model in this dataset is lower than CSLCM model, but the contour detection effect of some images is better CSLCM model. Since the proposed algorithm produces soft contour maps, a choice of scale is required for $R(x, y)$ that need obtain a single segmentation output. Similar to the performance measurement of the gPb algorithm, one possibility for scale choice is to use a fixed threshold (calibrated on a training set to provide optimal performance) for all images in the dataset. This is defined as the optimal dataset scale (ODS). The performance was also evaluated for an optimal threshold selected by an oracle on a per-image basis. It is possible to obtain better segmentations using this optimal image scale (OIS). We can see that some machine-learning based models, e.g., the *Probability of Boundary (Pb)* [39] and *Global Pb*
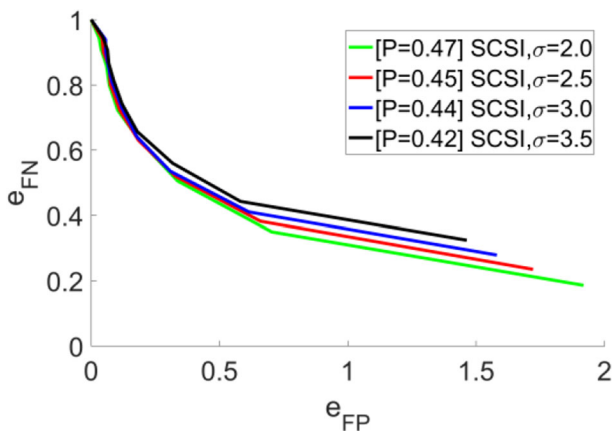

Fig. 10 The $e_{FP} - e_{FN}$ curve of proposed model with parameter $\sigma = [2.0, 2.5, 3.0, 3.5]$, respectively

**Table 1** Quantitative comparison of various models on the RuG grayscale images with P-Measure

Model	RuG40		
	ODS	OIS	Best P-scores
Biologically -motivated			
SCSI	0.47	0.52	0.54
CSLCM [11]	0.49	0.51	0.56
MCI [56]	0.47	0.49	0.52
Butterfly-shaped [61]	0.45	0.48	0.49
ISO-gd [56]	0.41	0.44	0.47
Computational / Machine-learning			
Pb (BG+TG) [39]	0.46	0.50	0.50
GPb [3]	0.48	0.56	0.56

(gPb) [3] did not perform well across different dataset, because the biologically-motivated model simulates the information processes mechanism of human visual system, and usually uses low-level image features, which is not dataset-dependant.

3.2.2 BSDS dataset

The performance of the proposed model was further evaluated using a publicly available dataset, the Berkeley Segmentation Data Set (BSDS300 and BSDS500). Each image in the dataset has multiple human-labeled segmentations as the ground truth data. The so-called F-measure [39] was also computed, as below:

$$F = \frac{2*Precision*Recall}{Precision + Recall},$$ (22)
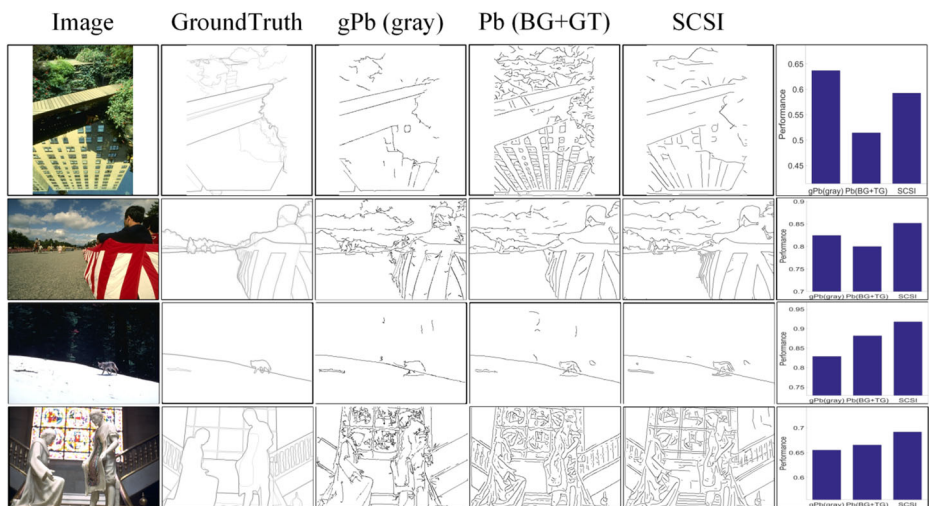
Fig. 11 Results of various models on four example images from the BSDS300 dataset. The columns from left to right represent the original image, corresponding ground truth, *gPb*, *Pb* detector using BG and TG cues, *SCSI* and corresponding performance (F- Score) obtained from its right contour map using the optimal parameter setting, respectively
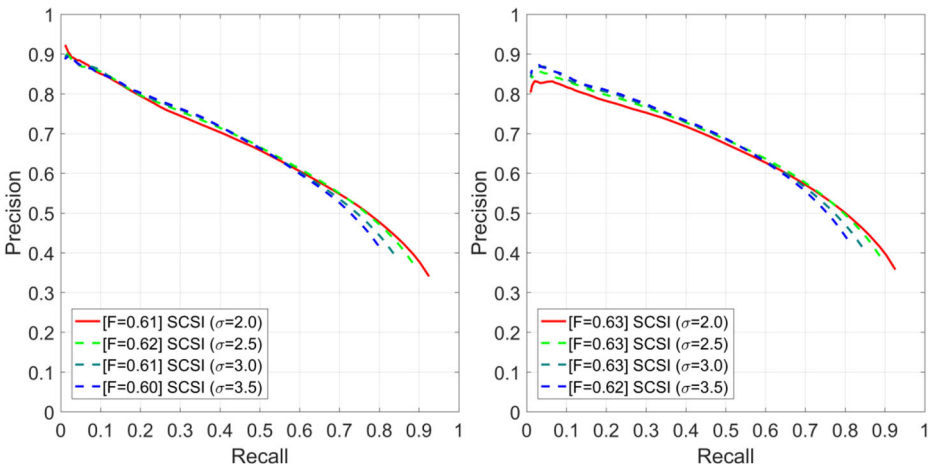
Fig. 12 The *Precision-Recall* curves of the *SCSI* contour detection models on the BSDS300 (left) and 500 (right) dataset

where *Recall* denotes the recall ratio reflective of the probability that the detected edge is valid and *Precision* denotes the precision ratio reflective of the probability that the ground truth edge was detected.

Figure 11 shows four images and the binary boundaries detected by the *SCSI* model and two machine learning based algorithms: the *Pb* and *gPb (gray)*. As shown, the proposed non-learning methods have similar or even better *F*-Scores than these supervised learning contour detection methods. It can also be observed that the proposed model is better than *Pb* and *gPb* at some aspect of texture suppression in the four test images.

Quantitative performance measurement was also carried out using a test set of 100 and 200 images from BSDS300 and BSDS500, respectively. Figure 12 show the *Precision-Recall curve* [42] of the *SCSI* contour detection model on the BSDS300 and 500 datasets,
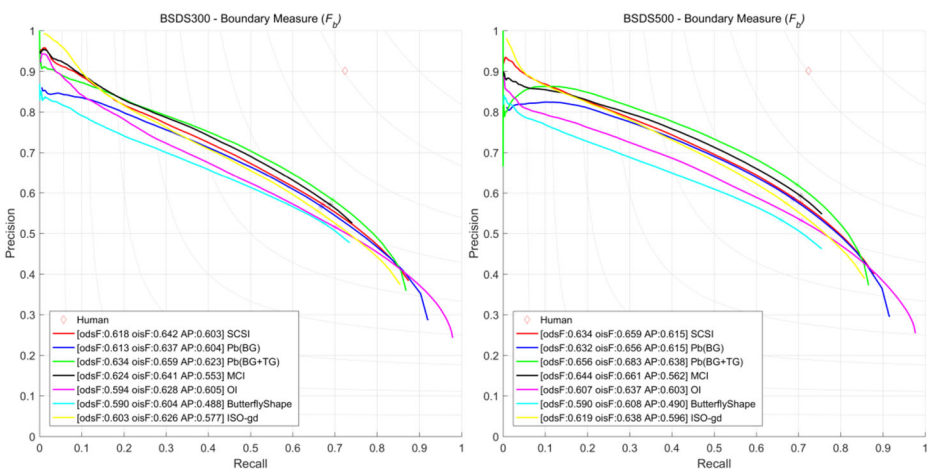
Fig. 13 Quantitative comparison of various models on the whole BSDS300 and BSDS500 dataset using the *Precision-Recall* curve. The legend shows the best ODS, OIS and AP on different contour detection algorithms

Table 2 Quantitative comparison of various models on the BSDS300 and BSDS500

Method	BSDS300			BSDS500		
	ODS	OIS	AP[1]	ODS	OIS	AP
Computational / Machine-learning						
MCG [4]	0.72	0.74	0.73	0.75	0.78	0.76
gPb [3]	0.69	0.70	0.73	0.70	0.74	0.71
BEL [20]	0.64	–	–	–	–	–
Mean Shift [15]	0.63	0.66	0.54	0.64	0.68	0.56
N-cuts [16]	0.62	0.66	0.43	0.64	0.68	0.45
Pb (BG+TG) [39]	0.63	0.66	0.62	–	–	–
Pb (BG) [39]	0.61	0.64	0.60	–	–	–
Canny [10]	0.58	0.62	0.58	0.60	0.63	0.58
Biologically -motivated						
SCSI	**0.62**	**0.64**	**0.61**	0.63	**0.66**	**0.61**
CSLCM [11]	0.61	**0.64**	0.60	0.62	**0.66**	**0.61**
MCI [56]	**0.62**	**0.64**	0.54	**0.64**	**0.66**	0.56
OI [56]	0.59	0.63	0.60	0.60	0.63	0.59
CO [55]	0.60	0.63	0.60	0.61	0.64	0.61
PC/BC-V1+lateral+texture [47]	0.61	0.63	0.40	**0.64**	0.65	0.41
ISO-gd [56]	0.60	0.63	0.59	–	–	–
Butterfly-shaped [61]	0.59	0.61	0.49	–	–	–

respectively. It can be seen that *SCSI* model show good performance when the scale parameter $\sigma = 2.5$.

In Fig. 13, the *Precision-Recall* curve [42] based comparison is given for our *SCSI* model ($\sigma = 2.5$) relative to the *MCI, OI, Pb (BG), Pb (BG + GT), Butterfly-shaped* and *ISO-gd* on the BSDS300 and 500 dataset. These graph shows that the proposed model outperforms the *OI, Pb (BG), Butterfly-shaped* and *ISO-gd*, similar to *MCI* but remains inferior to the *Pb (BG + GT)*. However, the *Pb* method need extra supervised and unsupervised learning processing with additional cues, such as textures; the *MCI* model uses not only orientation but also luminance and luminance contrast. Even with that, the proposed model has the same performance as the *MCI* and slightly lower than *Pb (BG + GT)* contour detectors.

Table 2 show quantitative comparisons from the BSDS300 and BSDS500, the optimal thresholds for both of them are 0.03. In this dataset our model is better than CSLCM model. As can be observed, the proposed *SCSI* operator receives a score of 0.62 on the BSDS300, higher than most of the other biologically-motivated methods. The scores of these methods remain below those of most learning-based algorithms; however, the proposed non-learning method still outperforms the *Pb (BG)* ($F = 0.60$).

4 Discussion and conclusion

It is widely believed that the primary visual cortex is involved in edge detection. In this work, a novel contour detection model inspired by the information processing mechanisms of the neural CRF and non-CRF responses in V1 was proposed. A subfield combination framework was utilized in place of the traditional pixel based surround inhibition structure. Two inhibition cues were employed on the basis of this framework: directional saliency and energy distribution. The proposed framework also makes use of two combination methods, center-surround and surround-surround inhibition.

Some drawbacks to the architecture of surrounding region, such as main problem requiring consideration was the size of the subfield. One limitation of the proposed strategy is the reliance it has on heuristics rather than general principles. For example, the simple tests performed in this study showed that the contour detection performance will be better if the framework uses a greater number of subfields of smaller size. However, excessively increasing the number of subfields not only increases the required computational power significantly, but also reduces the performance of salient contour extraction. Overlapping subfields also extract numerous useless features that cause imprecision in the degree of inhibition. However, too small a quantity and size of subfield is selected (such that the subfield-based method is degraded as in pixel-based frameworks), the performance degrades even from versions using excessive subfields. Therefore, balancing the number and size of subfields is something of an empirical decision. The choice made in this manuscript was primarily motivated by conceptual and experimental simplicity, while the theoretical study of optimal solutions was left for future research.

Aside from the structure of the subfield, performance also relied upon on efficient feature or cue extraction. In the proposed model, energy distribution cues were first employed as inspired by [56], in which the local energy distribution computation is sensitive to noise, and incorrect orientation estimates will result in ineffective surround inhibition. However, this experiment found that the disorder of the region itself also can be obtained as a useful clue for texture inhibition. Furthermore, surround-surround interactions were creatively employed to compute the inhibition term. Given this, the inhibition term does not accept information from the central region, but is able to efficiently suppress some isolated edges surrounded by cluttered backgrounds.

Finally, it is necessary to comment on some possible future improvements of the proposed models. Considering the importance of color information in visual pathways, one possible improvement direction is to introduce color features to our gray-scale-based model in two ways: (1) coding color features into one channel as pre-processing; (2) using color instead of gray features in subfield coding, i.e., transferring $\vec{E}(x, y)$ mentioned in (6) to a higher dimension vector. In addition, the color information could be exploited by using either individual R-G-B information or the features after single- and double-opponency processing [22, 55] or both.

In summary, this study proposed a subfield based model by combining center-surround and surround-surround inhibition effects to improve contour detection in cluttered scenes. With the displayed competitive performance in comparison to current state-of-the-art algorithms, the proposed subfield based framework is expected to facilitate efficient computational models in the field of machine vision.

References

1. Albright TD, Stoner GR (2002) Contextual influences on visual processing. Annu Rev Neurosci 25(1):339–379

2. Angelucci A, Bressloff PC (2006) Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. Prog Brain Res 154:93–120
3. Arbelaez P, Maire M, Fowlkes C, Malik J (2011) Contour detection and hierarchical image segmentation. IEEE Trans Pattern Anal Mach Intell 33(5):898–916. https://doi.org/10.1109/TPAMI.2010.161
4. Arbeláez P, Pont-Tuset J, Barron JT, Marques F, Malik J (2014) Multiscale combinatorial grouping. IEEE Conference on Computer Vision and Pattern Recognition:328–335
5. Azzopardi G, Petkov N (2012) A CORF computational model of a simple cell that relies on LGN input outperforms the Gabor function model. Biol Cybern 106:1–13
6. Bar M (2004) Visual objects in context. Nat Rev Neurosci 5(8):617–629
7. Bertasius G, Shi J, Torresani L (2015) Deepedge: a multi-scale bifurcated deep network for top-down contour detection. In: IEEE conference on computer vision and pattern recognition, pp 4380–4389
8. Boukerroui D, Noble JA, Brady M (2004) On the choice of band-pass quadrature filters. Journal of Mathematical Imaging and Vision 21(1–2):53–80
9. Bredfeldt CE, Ringach D (2002) Dynamics of spatial frequency tuning in macaque V1. J Neurosci 22(5): 1976–1984
10. Canny J (1986) A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 6:679–698. https://doi.org/10.1109/TPAMI.1986.4767851
11. Cao YJ, Lin C, Pan YJ, Zhao HJ (2019) Application of the center–surround mechanism to contour detection. Multimed Tools Appl 78(17):1–21
12. Cavanaugh JR, Bair W, Movshon JA (2002) Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. J Neurophysiol 88(5):2547–2556
13. Chao-Yi L, Wu L (1994) Extensive integration field beyond the classical receptive field of cat's striate cortical neurons—classification and tuning properties. Vis Res 34(18):2337–2355
14. Coen-Cagli R, Dayan P, Schwartz O (2012) Cortical surround interactions and perceptual salience via natural scene statistics. PLoS Comput Biol 8(3):e1002405
15. Comaniciu D, Meer P (2002) Mean shift: a robust approach toward feature space analysis. IEEE Trans Pattern Anal Mach Intell 24(5):603–619
16. Cour T, Benezit F, Shi J (2005) Spectral segmentation with multiscale graph decomposition. Computer vision and pattern recognition, 2005. In: IEEE computer society conference on, vol 2. CVPR, pp 1124–1131
17. Croner LJ, Kaplan E (1995) Receptive fields of P and M ganglion cells across the primate retina. Vis Res 35(1):7–24
18. Das A, Gilbert CD (1999) Topography of contextual modulations mediated by short-range interactions in primary visual cortex. Nature 399(6737):655–661
19. Daugman JG (1985) Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. JOSA A 2(7):1160–1169
20. Dollar P, Tu Z, Belongie S (2006) Supervised learning of edges and object boundaries. CVPR 2:1964–1971
21. Duda RO, Hart PE (1973) Pattern classification and scene analysis. Wiley, New York
22. Gao S, Yang K, Li C, Li Y (2013) A color constancy model with double-opponency mechanisms. ICCV: 929–936
23. Girshick R (2015) Fast r-cnn. In: International Comference on computer vision, pp 1440–1448
24. Grigorescu C, Petkov N, Westenberg MA (2003) Contour detection based on nonclassical receptive field inhibition. IEEE Transactions on Image Processing 12(7):729–739. https://doi.org/10.1109/TIP.2003.814250
25. Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol 160(1):106–154
26. Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. J Physiol 195(1):215–243
27. Hummel RA, Zucker SW (1983) On the foundations of relaxation labeling processes. IEEE Trans Pattern Anal Mach Intell 3:267–287
28. Itti L, Koch C (2000) A saliency-based search mechanism for overt and covert shifts of visual attention. Vis Res 40(10):1489–1506
29. Jones JP, Palmer LA (1987) An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. J Neurophysiol 58(6):1233–1258
30. Kass M, Witkin A, Terzopoulos D (1988) Snakes: active contour models. Int J Comput Vis 1(4):321–331
31. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Aystems:1097–1105
32. Landy MS, Graham N (2004) 73 visual perception of texture. The Visual Neurosciences 1106

33. Li C-Y (1996) Integration fields beyond the classical receptive field: organization and functional properties. Physiology 11(4):181–186
34. 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 96(7):4052–4056
35. Lin C, Xu G, Cao Y, Liang C, Li Y (2016) Improved contour detection model with spatial summation properties based on nonclassical receptive field. Journal of Electronic Imaging 25(4):043018–043018
36. Maninis K-K, Pont-Tuset J, Arbeláez P, Van Gool L (2017) Convolutional oriented boundaries: from image segmentation to high-level tasks. arXiv preprint arXiv:1701.04658 40 (4): 819-833. https://doi.org/10.1109/TPAMI.2017.2700300
37. Marr D, Hildreth E (1980) Theory of edge detection. Proc R Soc Lond B Biol Sci 207(1167):187–217
38. Martin D, Fowlkes C, Tal D, and Malik J (2001) A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE international conference on, vol. 2, pp. 416–423.
39. Martin DR, Fowlkes CC, Malik J (2004) Learning to detect natural image boundaries using local brightness, color, and texture cues. IEEE Trans Pattern Anal Mach Intell 26(5):530–549. https://doi.org/10.1109/TPAMI.2004.1273918
40. Papari G, Petkov N (2011) Edge and line oriented contour detection: state of the art. Image Vis Comput 29(2):79–103
41. Papari G, Petkov N (2011) An improved model for surround suppression by steerable filters and multilevel inhibition with application to contour detection. Pattern Recogn 44(9):1999–2007
42. Pont-Tuset J, Marques F (2016) Supervised evaluation of image segmentation and object proposal techniques. IEEE Trans Pattern Anal Mach Intell 38(7):1465–1478
43. Prewitt JM (1970) Object enhancement and extraction. Picture Processing and Psychopictorics 10(1):15–19
44. Series P, Lorenceau J, Frégnac Y (2003) The "silent" surround of V1 receptive fields: theory and experiments. Journal of Physiology-Paris 97(4):453–474
45. Shen ZM, Xu WF, Li CY (2007) Cue-invariant detection of Centre–surround discontinuity by V1 neurons in awake macaque monkey. J Physiol 583(2):581–592
46. Shen W, Wang X, Wang Y, Bai X, Zhang Z (2015) Deepcontour: a deep convolutional feature learned by positive-sharing loss for contour detection. In: IEEE conference on computer vision and pattern recognition, pp 3982–3991
47. Spratling MW (2013) Image segmentation using a sparse coding model of cortical area V1. IEEE Trans Image Process 22(4):1631–1643
48. Tang Q, Sang N, Zhang T (2007) Extraction of salient contours from cluttered scenes. Pattern Recogn 40(11):3100–3109
49. Tang Q, Sang N, Liu H (2016) Contrast-dependent surround suppression models for contour detection. Pattern Recogn 60:51–61
50. Walker GA, Ohzawa I, Freeman RD (2000) Suppression outside the classical cortical receptive field. Vis Neurosci 17(03):369–379
51. Wei H, Lang B, Zuo Q (2013) Contour detection model with multi-scale integration based on non-classical receptive field. Neurocomputing 103:247–262
52. Wei H, Dai Z-L, Zuo Q-S (2016) A ganglion-cell-based primary image representation method and its contribution to object recognition. Connect Sci 28(4):311–331
53. Xie S, Tu Z (2015) Holistically-nested edge detection. In: International comference on computer vision, pp 1395–1403
54. 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
55. Yang K, Gao S, Li C, Li Y (2013) Efficient color boundary detection with color-opponent mechanisms. In: IEEE Conference on Computer Vision and Pattern Recognition, pp 2810–2817
56. Yang K-F, Li C-Y, Li Y-J (2014) Multifeature-based surround inhibition improves contour detection in natural images. IEEE Trans Image Process 23(12):5020–5032
57. Yang K-F, Gao S-B, Guo C-F, Li C-Y, Li Y-J (2015) Boundary detection using double-opponency and spatial sparseness constraint. IEEE Trans Image Process 24(8):2565–2578
58. Yang K-F, Gao S-B, Li Y-J (2015) Efficient illuminant estimation for color constancy using grey pixels. CVPR:2254–2263
59. Yang K-F, Li H, Li C-Y, Li Y-J (2016) A unified framework for salient structure detection by contour-guided visual search. IEEE Trans Image Process 25(8):3475–3488
60. Zeng C, Li Y, Li C (2011) Center–surround interaction with adaptive inhibition: a computational model for contour detection. NeuroImage 55(1):49–66
61. Zeng C, Li Y, Yang K, Li C (2011) Contour detection based on a non-classical receptive field model with butterfly-shaped inhibition subregions. Neurocomputing 74(10):1527–1534