



Application of the center–surround mechanism to contour detection

Yi-Jun Cao¹ · Chuan Lin¹  · Yi-Jian Pan¹ · Hao-Jun Zhao¹

Received: 5 October 2018 / Revised: 9 April 2019 / Accepted: 1 May 2019 /
Published online: 21 May 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Physiological studies have revealed that the center–surround mechanism widely exists in the primary stages of the human visual system, such as the retina, lateral geniculate nucleus (LGN), and primary visual cortex (V1). In retina ganglion cells (RGC) and the LGN, the mechanism is well known to have two types: center “on” and center “off.” However, this mechanism in V1 is shown as classical receptive field (CRF) stimulation and surrounding non-CRF suppression. Although these two manifestations differ in function and appear in different areas of the visual pathway, from the perspective of computational simulation, they simply compute the differences between the center and its surrounding information. In the past decade, many bio-inspired computational models have demonstrated that the center–surround mechanism is good at extracting salient contours while suppressing textures. Based on this mechanism, we propose a method for extracting local center–surround contrast information from nature images by using a normalized difference of Gaussian (DoG) function and a sigmoid activated function. Compared with previous contour detection models (especially bio-motivated ones), the proposed method can efficiently suppress textures more quickly and accurately. More importantly, the proposed algorithm yields even better contour detection, yet the computational complexity is similar to the classical Canny operator.

Keywords Center-surround mechanism · Surround suppression · Contrast information · Contour detection

1 Introduction

Contour extraction is a fundamental task in the field of computer vision and image processing [30]. A contour represents a change of pixel ownership from one object to another; in

✉ Chuan Lin
chuanlin@gxust.edu.cn

¹ College of Electric and Information Engineering, Guangxi University of Science and Technology, Donghuan Road 268, Liuzhou, Guangxi, China

comparison, edges focus on abrupt changes in low-level image features, such as brightness or color [28]. One approach for increasing the performance of contour detection involves integrating the local texture information into the coherence features and extracting efficient features for texture inhibition. 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.

Numerous visual physiological studies have shown that the primary visual cortex (V1) plays an important role in the perception of target contour information from visual scenes [1, 7, 12, 15–17, 20]. The pioneering work of Hubel and Wiesel [14, 15] revealed that the majority of neurons in V1 prioritize a small excitation area known as the classical receptive field (CRF), which is exquisitely sensitive to oriented bars or edges. Subsequently, more physiological findings have clearly indicated that a peripheral region known as the non-classical receptive field (non-CRF) exists around the CRF in most V1 neurons [20]. Although the separate stimuli of the non-CRF do not affect overall neuron responses, they can significantly modulate, sometimes suppress, the response of V1 in the CRF [1]. This CRF stimulation and non-CRF suppression mechanism is known as the center–surround interaction in V1.

Based on the insight of the non-CRF inhibition mechanism, several non-CRF inhibition–based contour detection models have recently been proposed. Grigorescu et al. [13] used the Gabor function to simulate the response of the CRF, and utilized a larger ring-shaped function to capture information about the surrounding non-CRF. More specifically, their model uses the non-negative difference of Gaussians (DoG) function to take the distance weight between the center and surrounding non-CRF by convolution with CRF responses. Compared to traditional edge detection, this model can effectively enhance the detection accuracy of target contours in front of complex visual backgrounds. Virtually all bio-inspired methods have subsequently used it to suppress texture [4, 21–23, 31, 34–37, 40, 43, 44]. These methods only consider the CRF responses as inhibition cues and ignore the local luminance information. Thus, subsequent methods, such as multiple-cue inhibition (MCI)[40], exploit some new inhibition cues by using luminance contrast to compute the center–surround difference. These methods construct a local sliding window, taking the gray image as input, and compute the local center–surround difference to obtain the probability of textures.

However, the above luminance-contrast–based models have the following drawbacks: (1) they are not effective when luminance contrast is used as an independent inhibitor, because this inhibitor cannot capture contrast differences over a larger range, and (2) the non-convolutional computation process has a less efficient calculation speed. MCI [40] used the root-mean-square (RMS) to compute the difference between the center and its surrounding pixels in a small sliding window. Compared to the convolution operations optimized by most image processing toolboxes, separate sliding window calculations result in additional computational overhead.

In this paper, we propose a novel center–surround-based luminance contrast model (CSLCM) for contour detection that uses two simple yet effective methods to solve the two aforementioned problems. Our main contributions are as follows. (1) We use a normalized DoG function, which can capture the local center–surround difference of the brightness while being accelerated by a conventional convolution toolbox, to simulate the spatial characteristics of the retinal ganglion cells (RGCs) and the lateral geniculate nucleus (LGN). (2) We use a sigmoid activation function to make the difference more prominent in texture suppression. These two methods can greatly reduce the computation time compared to RMS while improving the texture suppression effect.

The remainder of this paper is organized as follows. After a short review of previous work in edge detection (Section 2), we describe CSLCM (Section 3). In Section 4, we evaluate the performance of CSLCM for the RuG40 dataset and two Berkeley segmentation datasets (BSDSs). Finally, Section 5 presents concluding remarks.

2 Works related to contour detection

In this section, we present a brief overview of the field of contour detection. We first describe the existing algorithms, which are divided into differential, local energy, machine learning, and bio-inspired methods. Then, we discuss the role of the DoG function in the previous contour detection models and the proposed model.

Differential methods Early edge detection methods usually made use of spatial differential operators through convolutions between filters and images. These methods computed spatial gradients along different orientations, making them easy to implement and fast to compute. One of the most well-known methods—the Canny edge-detection operator [6]—is a multi-level edge detection algorithm that sequentially (1) uses finite-difference estimates of first-order partial derivatives to calculate the magnitude and direction of the gradient, (2) thins the gradient amplitude through non-maximum suppression, and (3) connects edges with hysteresis thresholding. However, these methods did not consider contextual information or middle-/high-level information, which might promote suppression of useless textures.

Local energy methods Inspired by strong responses to highly ordered phase information, some contour detection methods focus on the analysis of local phase information [29] and phase congruency [19] by using advanced techniques such as inferential statistics [18, 28]. Local energy approaches perform, for practical applications, similarly to the faster and conceptually simpler differential methods, although they only partially achieve their ability.

Machine learning methods More recent approaches use machine learning techniques for cue combinations. Martin et al. [28] proposed the Pb method, which considers multiple cues (including color, brightness, and texture) as the input of the logical regression classifier to extract and localize boundaries. Dollar et al. [11] proposed a supervised learning algorithm, called boosted edge learning (BEL), that attempts to learn edge classifiers in the form of a probabilistic boosting trees from thousands of simple image features. An advantage of this approach is that it may be possible to handle cues such as parallelism and completion in the initial classification stage. Mairal et al. [26] created both generic and class-specific edge detectors by learning from discriminative sparse representations of local image patches and further proposed a multi-scale method to extend sparse signal models to feature selection. Other Pb-based algorithms improve the capability of boundary detection by adopting multi-scale mechanisms or global information [3, 33]. For example, mPb considers local boundary cues, including contrast, localization, and relative contrast, and trains a classifier to integrate them across scales [33], and gPb integrates multiple local cues into a global framework based on spectral clustering [3]. Such algorithms usually integrate multi-scale information with a supervised learning method. Some other pattern recognition or image segmentation methods based on machine learning include mean shift [9], which provides a clustering framework, and the normalized cut (N-

cut) algorithm [10], which is a multi-scale spectral algorithm that uses the normalized-cut graph-partitioning framework for image segmentation. In addition, Aràndiga et al. [2] proposed a neural-network-based edge detection algorithm that is insensitive to changes in illumination. The machine learning methods are dataset dependent. Their performance, which is reliant on training images and corresponding annotation labels, surpasses that of most non-learning methods in almost all public datasets. However, it decreases when the well-trained machine learning models are tested on a new dataset.

Bio-inspired methods Following another line of inquiry, there is a long history of employing early visual mechanisms for image analysis. Some researchers focused on local statistics patterns [8], involving contrast dependence, orientation tuning, and spatial asymmetry, and additionally employed other concepts used in visual applications, such as contour detection [39, 41]. In recent decades, many biologically motivated contour detection models have been proposed, and they show good performance for gray-scale natural images [4, 13, 21–23, 31, 34–37, 40, 43, 44]. Grigorescu et al. [13] 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 surrounding region, creating texture suppression with a distance cue. Following this framework, several improved models have been proposed. Tang et al. [35] proposed a suppression model based on the side inhibition region according to the cyclic inhibition characteristics of V1, which eliminates the local directed edges generated by a large number of complex textures in the background. In order to balance the side inhibition mechanism and weak contour regions, Zeng et al. [43, 44] introduced a butterfly-shaped inhibition model and added scale information into the calculation of inhibition weight, thereby preserving the weaker contour regions and further strengthening the integrity of the extracted contours. According to the horizontal interactions in V1, Xiao and Cai [38] introduced contextual influences into the contour detection, specifically without separating regions of excitatory and inhibitory lateral connections. Yang et al. [40] modulated the final neuronal inhibition responses through the integration of contextual information and cues in the visual system, thereby enriching the inhibition term. Tang et al. [36] proposed a surround inhibition model that uses contrast to modulate the inhibition term and finally achieved good performance. Wei et al. [37] simulated the early biological neural visual mechanism and used the DoG and 3-G models to adaptively express the image. Although bio-inspired methods have made huge progress in recent years, their performance is below that of machine learning methods and their computational complexity is greater than that of differential methods. This paper proposes an efficient bio-inspired contour detection algorithm.

Role of the DoG function As a differential operator, the DoG is able to capture local center-surround differences. Early DoG functions were treated as mathematical approximations of Laplace and Gaussian (LOG). Unlike derivatives of the Gaussian function, which can directly detect edges from an image (Fig. 1a), DoG is used to extract the responses near the edges, as shown in Figs. 1b–c. Grigorescu et al. [13] proposed a novel method that uses non-negative DoG to suppress texture by convolution with an edge map, as shown in Fig. 1d. Virtually all bio-inspired methods have subsequently used it to suppress texture [4, 21–23, 31, 34–37, 40, 43, 44]. According to the differential structure,

we use the DoG function, followed by absolute value response and sigmoid activation, to obtain luminance contrast features, as illustrated in Fig. 1e. Compared to previous bio-inspired models, which treat local density edges as texture, we argue that the difference of center-surround luminance contrast can better distinguish contours and textures.

3 Contour detection model

In this section, we propose a new contour detection model based on the center-surround interaction mechanism in V1. First, we simulate responses in the CRF of V1 neurons. Then, we describe the detailed procedure for extracting the new inhibition term. Finally, we calculate the final response of the target contour by subtracting the surround suppression from the CRF response. The general framework of CSLCM is shown in Fig. 2.

3.1 CRF response of V1 neurons

For the orientation-selective V1 neurons modeled in this study, we use the derivative of the two-dimensional (2-D) Gaussian function to describe their properties when responding to the stimuli within the CRF. The derivative of 2-D Gaussian function can be expressed as follows:

$$CRF(x, y; \theta, \sigma) = \frac{\partial g(\tilde{x}, \tilde{y}; \theta, \sigma)}{\partial \tilde{x}}, \tag{1}$$

$$g(\tilde{x}, \tilde{y}; \theta, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\tilde{x}^2 + \gamma^2 \tilde{y}^2}{2\sigma^2}\right), \tag{2}$$

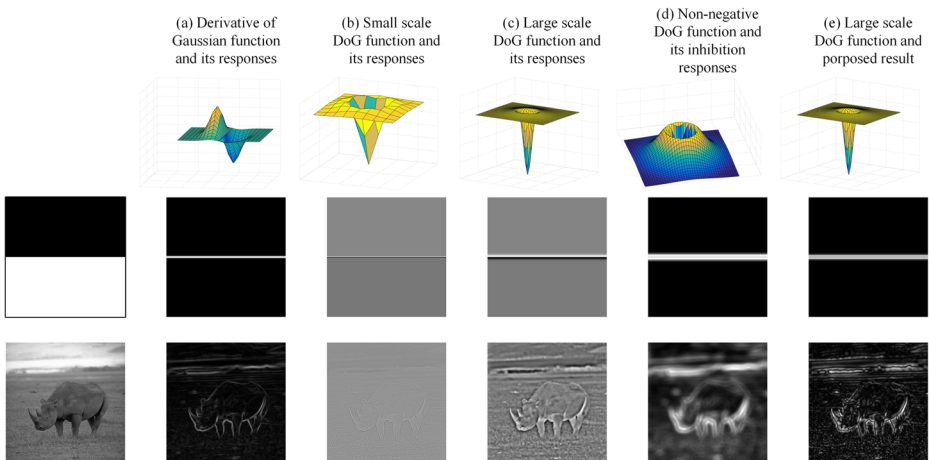


Fig. 1 Differential operators and their response maps on a synthetic image and a nature image. From top to bottom, (a-c) the functions and corresponding responses with two images; (d) the non-negative DoG function, the convolutional responses with edge map; (e) DoG function and proposed inhibition term

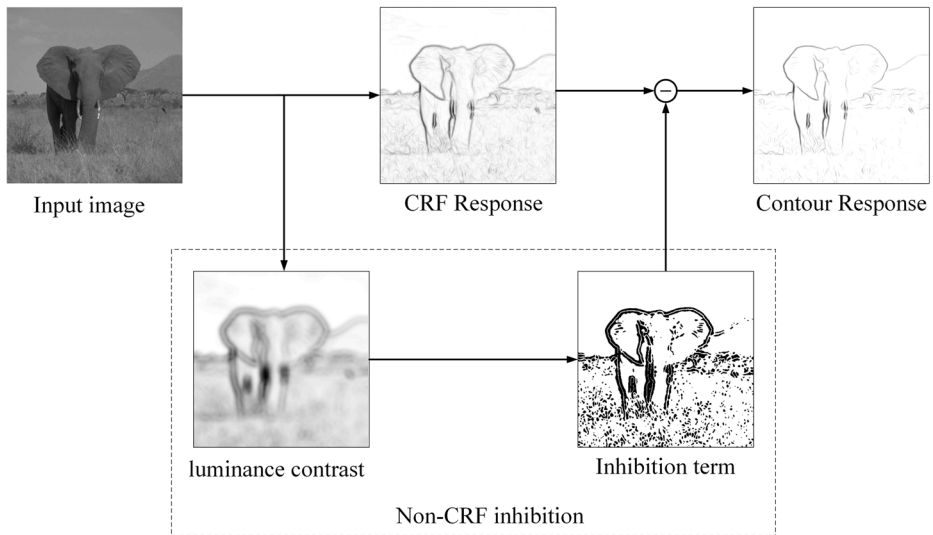


Fig. 2 General framework of CSLCM: the final response at each pixel is computed by subtracting the non-CRF inhibition from the CRF response

where $\tilde{x} = x\cos(\theta) + y\sin(\theta)$, $\tilde{y} = -x\sin(\theta) + y\cos(\theta)$. (x, y) and (\tilde{x}, \tilde{y}) are the original and rotated spatial coordinates, respectively. θ is the rotation angle (called the preferred orientation) of a neuron. The spatial aspect ratio, γ , and the standard deviation, σ , respectively determine the ellipticity and the size of CRF. In this paper, we set γ to 0.5 on the basis of physiological finding [32].

At each location, a pool of neurons with N_θ different preferred orientations, θ_i , is employed to process the local stimuli:

$$\theta_i = \frac{(i-1)\pi}{N_\theta}, i = 1, 2, \dots, N_\theta. \tag{3}$$

For an input image, $I(x, y)$, the CRF response of a V1 neuron with preferred orientation θ_i is computed as follows:

$$e_i(x, y; \theta_i, \sigma) = |I(x, y) * CRF(x, y; \theta_i, \sigma)|, \tag{4}$$

where $*$ denotes the convolution operation. The maximum CRF response over all N_θ neurons is selected as the final CRF response:

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

The orientation corresponding to the maximum CRF response, $\Theta^l(x, y)$, is computed as

$$\Theta^l(x, y) = \theta_k, \tag{6}$$

$$k = \operatorname{argmax}\{E_{\lambda, \sigma, \theta_i}(x, y) | i = 1, \dots, N_\theta\}. \tag{7}$$

3.2 Non-CRF inhibition responses

An experiment about the behavior of inhibition terms on synthetic images has shown that the gradient magnitude is strong on both isolated edges and textures [31]. Therefore, $CRF(x, y; \theta, \sigma)$ alone is not sufficient to discriminate between the two patterns. In contrast, the inhibition term is much higher on textures than on isolated edges. Thus, when the inhibition term is subtracted from the gradient magnitude, the resulting quantity has a strong response on isolated edges only.

The local center–surround luminance contrast is an important visual feature for understanding natural scenes. Many studies have reported the statistics of local luminance and contrast in natural images, but with inconsistent conclusions. For example, Mante et al. [27] claimed that local luminance and luminance contrast are independent (or weakly correlated) in early visual systems and in natural scenes. However, Lindgren et al. [24] revealed a strong spatial correlation between local luminance and luminance contrast in natural images.

In this work, we evaluate the contribution of local luminance contrast in contour detection with surround inhibition. Specifically, a normalized DoG function is used as the extraction operator:

$$w(x, y) = \begin{cases} \frac{DoG(x, y; \sigma, k)}{\sum_{DoG(x, y; \sigma, k) < 0} DoG(x, y; \sigma, k)}, DoG(x, y; \sigma, k) < 0, \\ \frac{DoG(x, y; \sigma, k)}{\sum_{DoG(x, y; \sigma, k) \geq 0} DoG(x, y; \sigma, k)}, DoG(x, y; \sigma, k) \geq 0, \end{cases} \tag{8}$$

$$DoG(x, y; \sigma, k) = \frac{1}{2\pi(k\sigma)^2} e^{-\frac{x^2+y^2}{2(k\sigma)^2}} - \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \tag{9}$$

where k represents the ratio of two different scales. Then, the response local luminance contrast $LC(x, y)$ is computed by convolution:

$$LC(x, y) = |I(x, y) * w(x, y)|. \tag{10}$$

Intuitively, it is rational to suppose that if the luminance contrast obeys $LC(x, y) \leq T$, then the difference between the brightness of the center and that of its surroundings is small in this local range, and vice versa. Thus, a sigmoid function is used to activate the $LC(x, y)$, and the final inhibition response, with soft threshold T , is obtained:

$$Inh(x, y) = 1 - sigmoid(LC(x, y) - T, e^\beta), \tag{11}$$

$$sigmoid(x, z) = \frac{1}{1 + z^{-x}}, \tag{12}$$

where $sigmoid(x, z)$ is usually used as an activation function in a neural network, and β controls the smoothness of the $sigmoid$ function. Equations (11) and (12) show that $Inh(x, y)$ is near 1 or 0 when $LC(x, y)$ is less than or greater than parameter T , respectively. As β increases, $Inh(x, y)$ approaches the binary function; in Section 4.2.1, quantitative experiments reveal that the performance of the binary function is better than the ‘soft’ $sigmoid$ function.

3.3 Contour extraction

The final contour response is the combination of the CRF response, $E(x, y; \theta)$, and the inhibition term, $Inh(x, y)$:

$$r(x, y) = H(E(x, y; \theta) - \alpha Inh(x, y)), \quad (13)$$

$$H(z) = \begin{cases} 0, & z < 0, \\ z, & z \geq 0, \end{cases} \quad (14)$$

where the function $H(z)$ ensures that the operator yields non-negative responses and α is an inhibition factor that controls the connection strength between the CRF and non-CRF of neurons.

4 Results and discussion

In this section, we first provide a qualitative example of the inhibition effect of CSLCM. Then, we quantitatively measure the performance on the RuG dataset [13], which includes 40 natural gray-scale images with corresponding ground truths, and BSDS300 and BSDS500 [3, 28], which include 300 images and 5–10 associated human-marked segmentations for each image. Finally, we compare the computational complexity of CSLCM with other contour-detection models. Table 1 summarizes the meanings of the parameters involved in the proposed model. The parameter settings are also listed in Table 1, which are identical for both the RuG and BSDS300 datasets.

Using standard post-processing, the binary contour maps are constructed from RuG and BSDS by the standard procedure of non-maxima suppression (NMS) followed by hysteresis thresholding [13].

4.1 Qualitative experiments

In this section, we proceed through some qualitative and intuitive experiments comparing two kinds of center-surround inhibition methods: isotropic non-CRF inhibition model (ISO) [13], which constructs center-surround contrast by using edge probability cues, and center-surround luminance contrast cues [40]. It is worth noting that luminance contrast is only used

Table 1 Summary of parameter meanings and values

Parameter	Description	Equation	value
σ	Size of CRF and the standard deviation of Gaussian filter	(1)	[1.0,3.0]
γ	Spatial aspect ratio of the CRF filter	(2)	0.5
N_θ	Number of preferred orientations for neurons	(3)	8
k	Ratio of Gaussian functions with two different scales	(9)	[1.5,4]
T	Sigmoid function offset value along x -axis	(11)	[0.01,0.2]
β	Smoothness of the <i>sigmoid</i> function	(11)	[1300]
α	Coefficient to control the inhibition of surroundings	(13)	0.5

as one of the inhibition cues in MCI [40]. Thus, in this section, we also use luminance contrast inhibition (CI) for comparison [40].

4.1.1 Comparison with ISO model

As shown in Fig. 3b, much of the background texture remains in the CRF response map. In order to suppress textures, ISO uses the edge probability cues to obtain the inhibition term (Fig. 3d₁). In CSLCM, the inhibition term is constructed from the luminance contrast cues (Fig. 3c₁). Both inhibition terms target contours and background textures, but the luminance contrast cues are more efficient at textural inhibition, highlighting the difference between the contour and the background texture response. Figures 3c₂ and d₂ show the final responses for the target contour from CSLCM and the ISO model, respectively, which are calculated by subtracting the inhibition term from the CRF response. Figures 3c₃ and d₃ show the binary contour maps from CSLCM and the ISO model, respectively, which are constructed through the standard procedure of NMS followed by hysteresis thresholding. The binary contour map in Fig. 3c₃ is more complete and more effectively suppresses background textures than that in Fig. 3d₃. This contrast experiment sufficiently demonstrates that the extracted luminance contrast cues can further suppress background textures while more completely preserving the contour information.

4.1.2 Comparison with CI model

As with the previous example, many unwanted texture edges exist in the CRF response map (Fig. 4b). In order to suppress textures, the CI model [40] computes the inhibition term by

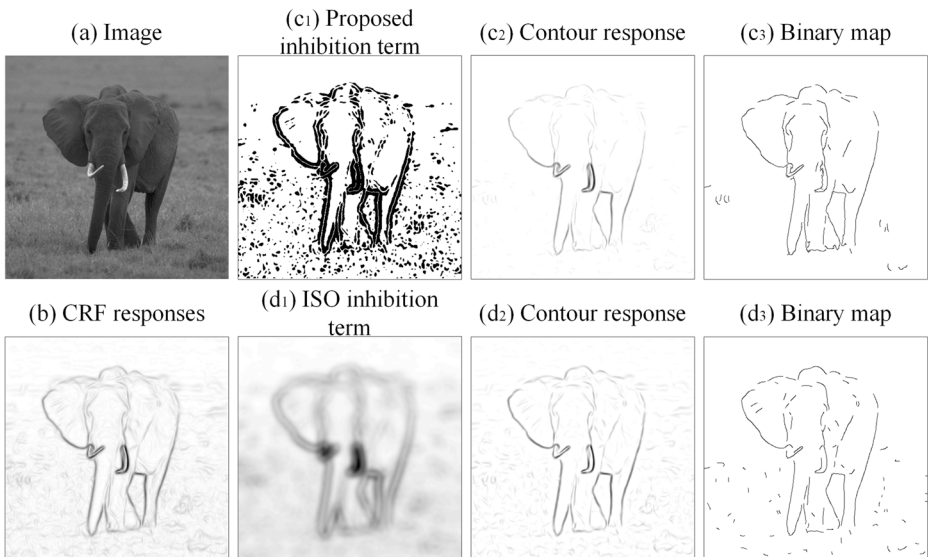


Fig. 3 Qualitative comparison between CSLCM and the ISO model: (a) input image, (b) gradient magnitude map as the neuronal responses to the stimuli within CRF, (c_{1–3}) the inhibition term $Inh(x, y)$ with $T = 0.05$, $\beta = 300$, final responses $r(x, y)$, and binary contour map, respectively, and (d_{1–3}) the inhibition term from the ISO model, final responses $r(x, y)$, and binary contour map, respectively

considering the luminance contrast cues and RMS calculations (Fig. 4d₁). With the limitation of the sliding window size and the lack of a corresponding activation function, the isotropic inhibition efficiently suppresses textures, but it also inhibits some perceptually salient contours, especially contours embedded in cluttered background. In the proposed model, the final inhibition term (Fig. 4c₁) is modulated by DoG and activated by a sigmoid function. Comparing Fig. 4c₃ and d₃ reveals that at object contours (e.g., the contour of bear), the inhibition strengths of CSLCM are generally weaker; in contrast, the texture regions (e.g., the water) receive relatively stronger inhibition with CSLCM. Consequently, our inhibition strategy is more efficient for texture suppression and contour protection. Hence, CSLCM can respond strongly to perceptually salient contours, but is relatively insensitive to textures as compared with the CI model [40].

4.2 Quantitative experiments

4.2.1 RuG dataset

We test CSLCM on the RuG dataset, which includes 40 natural images and associated ground truth binary contour maps drawn by hand. This dataset has been widely used to evaluate the performance of contour detectors [13, 21, 31, 35, 36, 40, 43, 44]. For quantitative evaluation of the performance of the proposed method, the binary contour maps will be constructed by the standard procedure of NMS followed by hysteresis thresholding [7]. We measure the similarity between the detected contour maps and corresponding known ground truth with the *P*-score [13]:

$$P = \frac{card(E)}{card(E) + card(E_{FP}) + card(E_{FN})}, \tag{15}$$

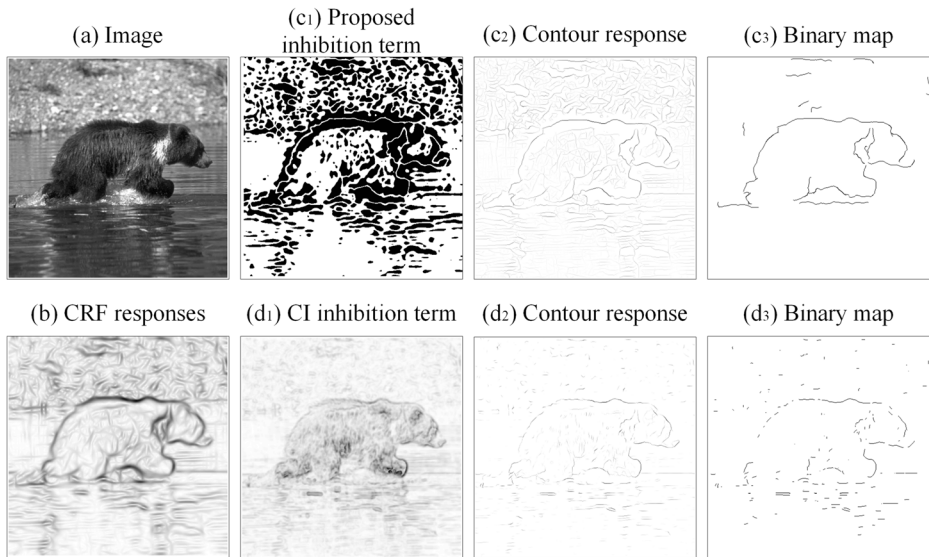


Fig. 4 Qualitative comparison between CSLCM and the CI model: **(a)** input image, **(b)** gradient magnitude map as the neuronal responses to the stimuli within CRF, **(c₁₋₃)** the inhibition term $Inh(x, y)$ with $T = 0.05$, $\beta = 300$, final responses $r(x, y)$, and binary contour map, respectively, and **(d₁₋₃)** the inhibition term from the CI model, final responses $r(x, y)$, and binary contour map, respectively

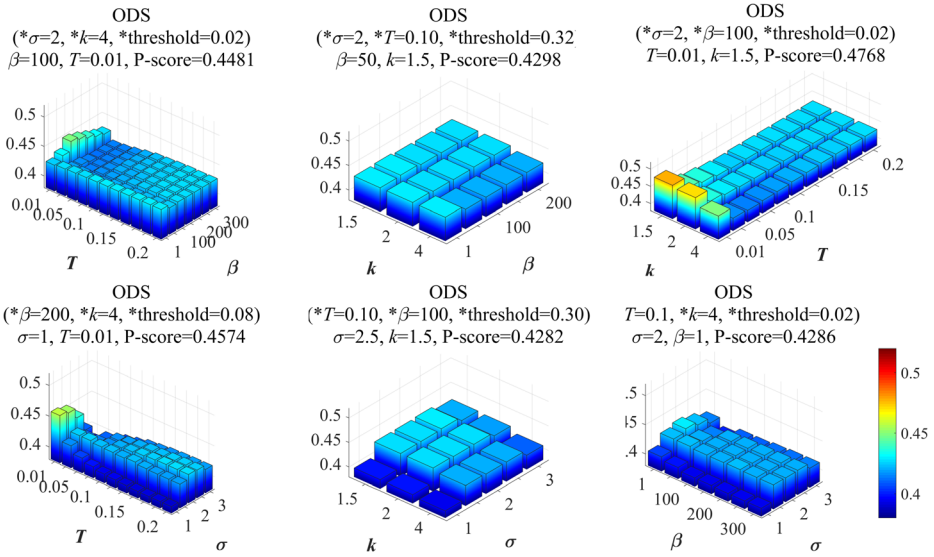


Fig. 5 Optimal dataset scale (ODS) of CSLCM over all images of the RuG dataset with various parameter combinations. Each bar in the figure represents the ODS over all the images for a given parameter combination. The text above each panel highlights the best ODS, which is obtained using the given parameter values. The symbol “*” indicates the fixed parameters

where $\text{card}(A)$ denotes the number of elements of set A , E refers to the set of contour pixels that are correctly detected by the algorithm, and E_{FP} and E_{FN} refer to the set of false positive

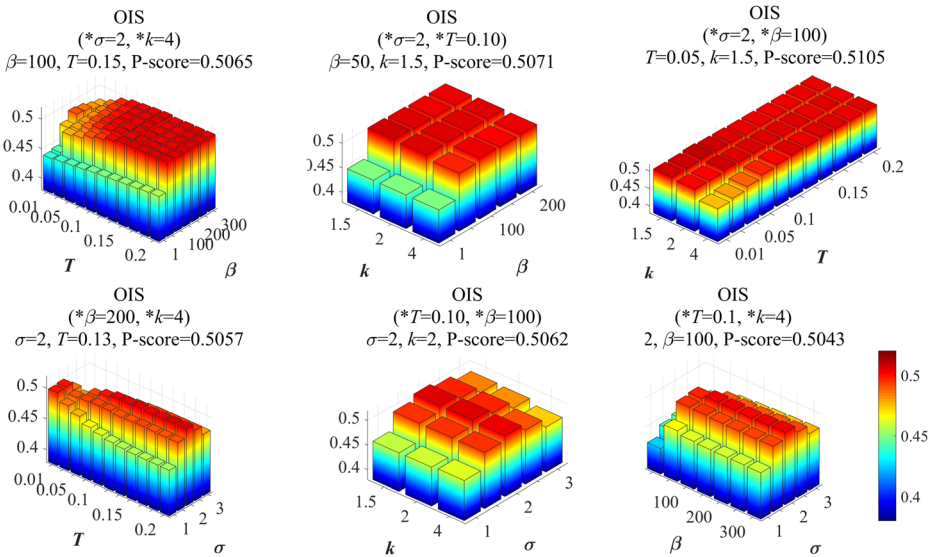


Fig. 6 Optimal image scale (OIS) of CSLCM over all images of the RuG dataset with various parameter combinations. Each bar in the figure represents the OIS over all the images for a given parameter combination. The text above each panel highlights the best OIS, which is obtained using the given parameter values. The symbol “*” indicates the fixed parameters

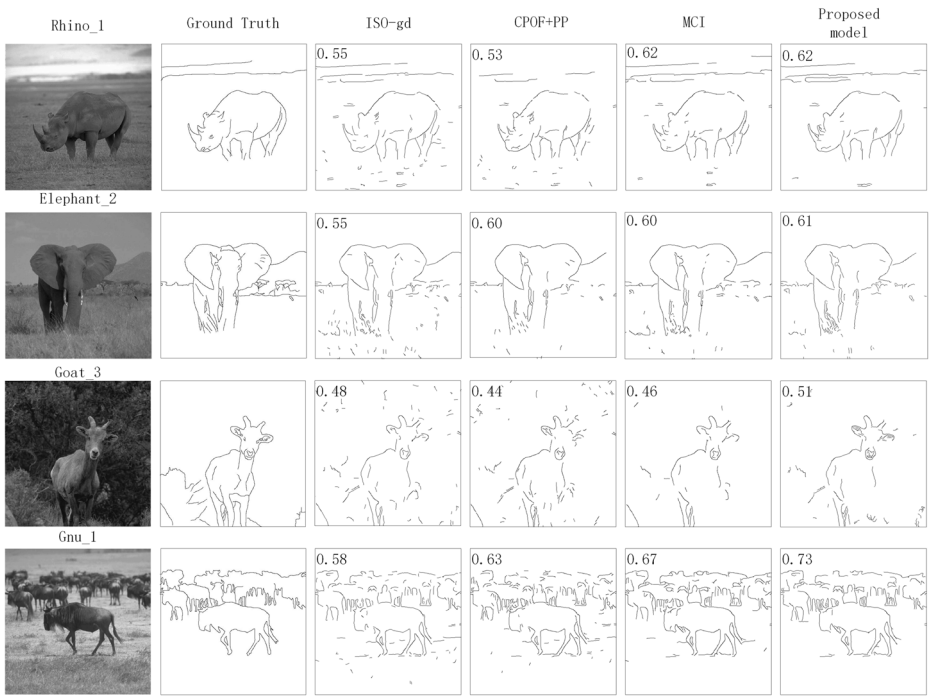


Fig. 7 Binary contour maps of four images from the RuG dataset: from left to right, the columns are input images, corresponding ground truth, and the ISO-gd [40], CORF-PP [5], MCI [40], and CSLCM models. The values in the upper-left corner are the P -scores

pixels and false negative pixels, respectively. The range of the P -score is [0, 1]; therefore, higher P -scores reflect better detection performance.

In this experiment, series of parameter combinations (5 values of σ within [1.0,3.0], 11 values of T within [0.01,0.2], 7 values of β within [1,300], and 3 values of k within [1.5, 4]) were selected for further statistical analysis. The parameter values are displayed in Table 1. In Figs. 5 and 6, the heights of the bars reflect, respectively, the optimal dataset scale (ODS) and optimal image scale (OIS) of P over all 40 images in the RuG dataset with

Table 2 Quantitative results for several contour detection algorithms on RuG40

Model	RuG40		
	ODS	OIS	Best P-scores
Biologically motivated			
CSLCM	0.49	0.51	0.56
MCI [40]	0.47	0.49	0.52
CORF-PP [5]	0.45	0.49	0.49
ISO-gd [40]	0.41	0.44	0.47
Computational/Machine learning			
Pb (BG + TG) [28]	0.46	0.50	0.50
gPb [3]	0.48	0.56	0.56

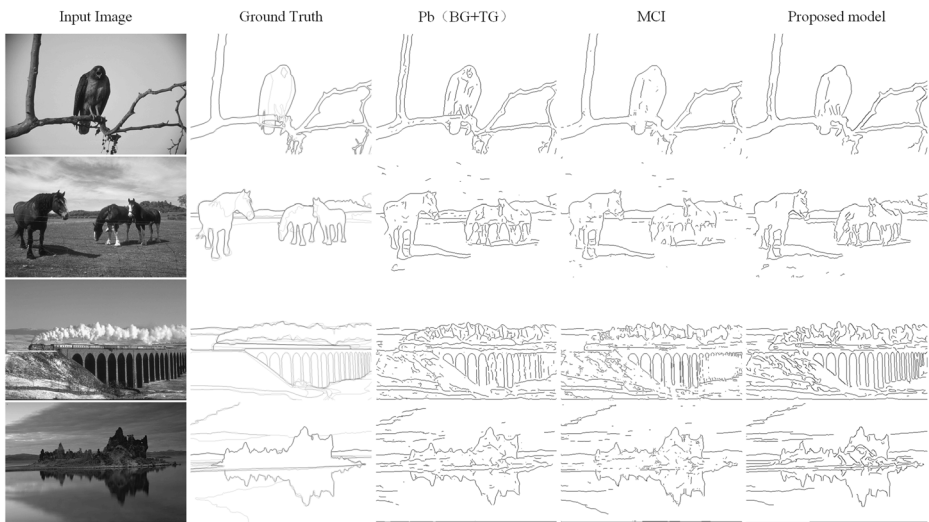


Fig. 8 Binary contour map on four images from the BSDS300 dataset. The columns from left to right are input images, corresponding ground truth, Pb (BG + TG) [28], MCI model [40], and CSLCM, respectively

different parameter combinations. Two parameters were 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. Comparing Figs. 5 and 6, we find an interesting fact in that some of the performance trends between ODS and OIS are opposite. For example, from the left two panels and the upper right panel of Fig. 5, the smaller σ , T , and k achieve better performance, whereas the corresponding panels of Fig. 6 exhibit that the larger T , k , and moderate σ achieve better performance. This may reveal a potential fact that in the case of a fixed threshold, a smaller receptive field size (small σ and k) can extract more stable local information, making the model more robust.

Intuitively, a “soft” inhibition term, e.g., $\beta = 1$, should yield more effective contour performance. However, in the above experiments, a “hard” inhibition term results in robust and good performance, as shown in the upper middle and bottom right panels of Figs. 5 and 6. This may be due to the fact that “hard” segmentation will cause strict texture inhibition or that a truncated activation response is more consistent with neuronal information transmission in the human visual system.

We evaluated and compared CSLCM with three other excellent biologically inspired models, i.e., ISO-gd [40], CORF-PP [5], and MCI [40], for four typical natural images in the RuG dataset, as shown in Fig. 7. Note that ISO-gd [40] uses the derivative of the 2D Gaussian function instead of the Gabor function for contour detection. The proposed method outperformed all the other models in terms of better suppression of responses to textures and better preservation of contours. The images show that CSLCM performs best in the trade-off between contour preservation and texture suppression. Specifically, the contour map obtained by CSLCM for Gnu_1, is very close to the ground truth, especially when compared with the other models.

We qualitatively measured the performance of some biologically motivated models, such as ISO-gd [40], CORF-PP [5], MCI [40], and CSLCM, and some computational/machine learning models, such as Pb (BG + TG) [28] and gPb [3], on the RuG dataset, as shown in Table 2. The best P-scores indicate that the parameters were optimized for each image and averaged over 40 images.

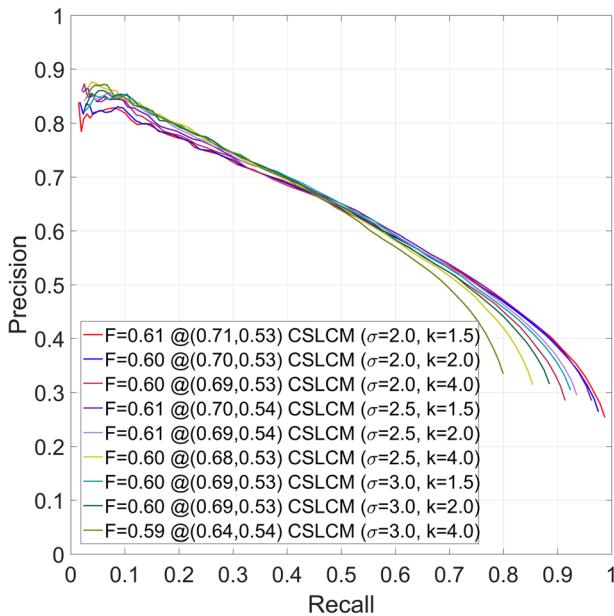


Fig. 9 Precision–Recall curves of CSLCM for BSDS300

CSLCM outperformed the compared models in the RuG dataset and reached a new state-of-the-art result (ODS = 0.49, OIS = 0.51, Best P-score = 0.56) compared with biologically motivated models for the RuG dataset. Note that we used the parameter combination ($\beta = 100$, $\sigma = 1.5$, $T = 0.01$ and $k = 1.5$) to measure the ODS and OIS performances. Interestingly, some machine learning based models, e.g., Pb [28] and gPb [3], did not perform well across different datasets because the biologically motivated models simulate the information processing mechanism of the human visual system, and usually use low-level image features, which are not dataset-dependent.

4.2.2 BSDS results

We further tested the proposed model with BSDS300 and BSDS500. Figure 8 presents several examples of the contour maps obtained with CSLCM followed by the NMS and binarization. The first and second columns of Fig. 8 show five test images from BSDS300 dataset and their human-marked boundaries, respectively. The third and fourth columns show the results of a machine learning based algorithm (Pb (BG + TG) [28]) and a state-of-the-art biologically motivated model (MCI [40]). The last column shows the results of CSLCM, showing that CSLCM is superior to the compared models in the continuity and preservation of contours.

The performance of CSLCM was further evaluated using BSDS300 and BSDS500. Each image in the dataset has multiple human-labeled segmentations as the ground truth data. The so-called F -score [28] was also computed for each contour:

$$F = \frac{2 * Precision * Recall}{Precision + Recall}, \quad (16)$$

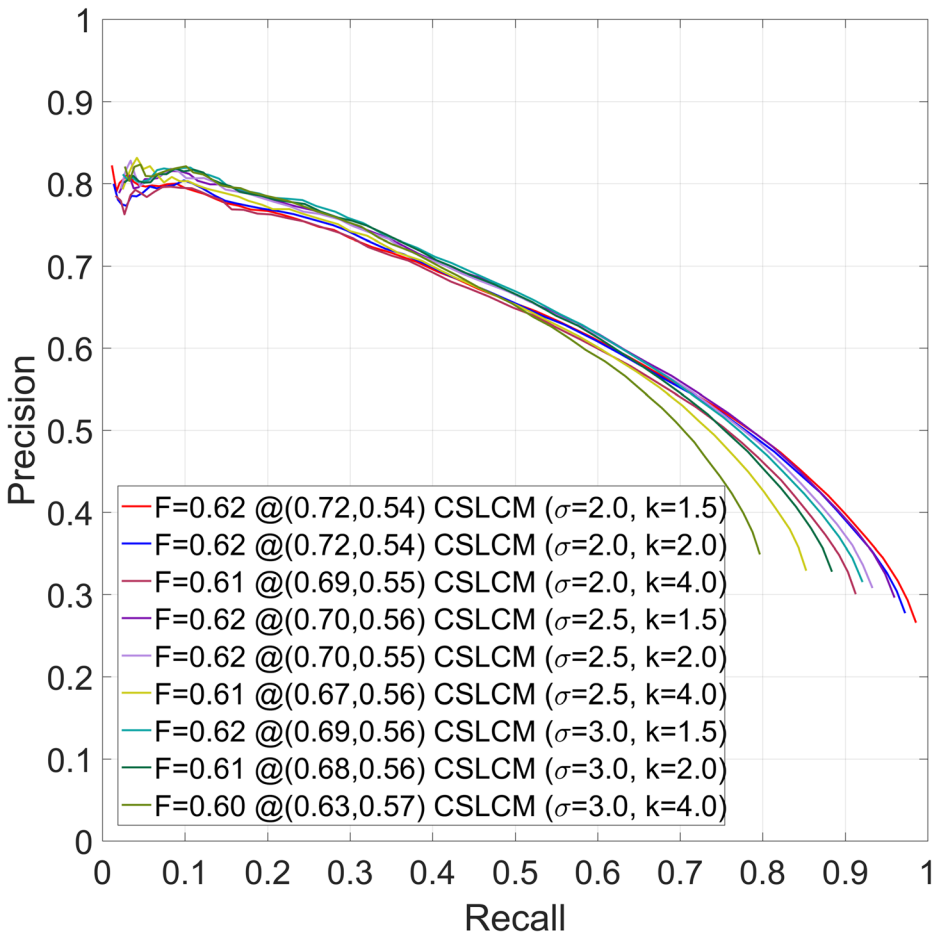


Fig. 10 Precision–Recall curves of CSLCM for BSDS500

where *Precision* denotes the fraction of detected edges that are valid and *Recall* denotes the fraction of ground truth edges that were detected. Note that we used tolerant parameters $\text{maxDist} = 0.0075$ in the BSDS benchmark.

Quantitative performance measurements were also performed using a test set of 100 images from BSDS300 and 200 images from BSDS500. Figures 9 and 10 show the Precision–Recall curves of CSLCM with different parameter combinations for BSDS300 and BSDS500, respectively. These results show that small k performs better than the value from experience ($k = 4$) in previous work [13, 22, 23, 25, 43, 44].

Table 3 shows the statistical results for BSDS300 and BSDS500. In this experiment, we set $\beta = 100$, $\sigma = 2.5$, $T = 0.05$, and $k = 1.5$. We can see that CSLCM has good performance compared with the other biologically motivated models for gray-scale images. The proposed method uses only center–surround luminance contrast for texture inhibition. The state-of-the-art method (MCI) used luminance, luminance contrast, and orientation cues for texture suppression. The model of PC/BC-V1 + lateral+texture [34] achieves its performance scores under the condition that recurrent lateral excitatory connections and other special lateral

Table 3 Quantitative results for several contour detection algorithms applied to the gray-scale images of BSDS300 and BSDS500

Method	BSDS300			BSDS500		
	ODS	OIS	AP ^a	ODS	OIS	AP
Computational/Machine learning						
gPb [3]	0.69	0.70	0.73	0.70	0.74	0.71
BEL [11]	0.64	–	–	–	–	–
Mean Shift [9]	0.63	0.66	0.54	0.64	0.68	0.56
N-cuts [10]	0.62	0.66	0.43	0.64	0.68	0.45
Pb (BG + TG) [28]	0.63	0.66	0.62	–	–	–
Pb (BG) [28]	0.61	0.64	0.60	–	–	–
Canny [6]	0.58	0.62	0.58	0.60	0.63	0.58
Biologically motivated						
CSLCM	0.61	0.64	0.60	0.62	0.66	0.61
MCI [40]	0.62	0.64	0.54	0.64	0.66	0.56
CI [40]	0.61	0.63	0.57	–	–	–
PC/BC-V1 + lateral+texture [34]	0.61	0.63	0.40	0.64	0.65	0.41
ISO-gd [40]	0.56	0.59	0.54	–	–	–
Butterfly-shaped [44]	0.59	0.61	0.49	–	–	–

^a AP: average precision

connections are incorporated into the classical PC/BC model of V1. CSLCM and the CI model [40] obtain similar ODS; however, CSLCM far exceeds CI [40] in calculation speed (see Section 4.3). Although CSLCM only used luminance contrast cues, the *F*-scores of OIS and AP for BSDS300 and BSDS500 reached state-of-the-art levels. Furthermore, the *F*-score of CSLCM outperforms some traditional machine learning–based methods (e.g., Pb (BG) [28]) for BSDS300.

4.3 Computational complexity

The proposed model has very low computational cost because its building blocks are simple convolutions with different kernels. With this in mind, we report the average computational time of some algorithms for BSDS500 in Table 4; although the MATLAB implementation of CSLCM is very slow, it outperforms nearly every method considered. We only used a single CPU core for computation and omitted the computational time of the post-processing actions of NMS and hysteresis thresholding.

Table 4 Average computational time (s) of seven-edge detection algorithms on BSDS500 under the MATLAB framework with Intel(R) Xeon(R) CPU E5–2650 v4 @ 2.20 GHz

Algorithm	Time (s)
Canny [6]	0.08
ISO [13]	0.21
MCI [40]	17.12
CI [40]	3.25
PC/BC-V1 + lateral+texture[34]	>> 1200
Pb (BG + TG) [28]	5.78
CSLCM	0.11

5 Conclusions

Since 1980, many psychophysical and neurophysiological studies have focused on the non-CRF in V1. Inspired by these studies, many computational models for contextual influences and center-surround interactions exist. Several of these models use various features for computing center-surround contrast to improve contour performance [13, 31, 40, 44], and others focus on simulating biological visual cellular mechanisms [42, 43]. In this paper, we focused on center-surround luminance contrast information and proposed a novel contour detection model. The proposed model performed well for three popular datasets: RuG, BSDS300, and BSDS500. More importantly, the proposed model is computationally very fast. The proposed CSLCM has computational complexity similar to Canny [6] but its performance is similar to that of MCI, which is better than Canny. It must be acknowledged that within the proposed framework, we have managed to only model a portion of what is known about center-surround interaction. The entire mechanism is considerably more complex. We simulated a simplified version of the center “off” function of the GCs and LGN layers to extract center-surround luminance contrast information. However, even in the early information processing truncation, i.e., the retina, the information processing of cells shows functional diversification and specificity, such as color-coding modeled cone cells [39, 41] and bright- and dim-coding modeled rod cells [45].

Biologically inspired solutions, such as the one presented here, make two contributions: technological (improving performance or reducing computational time) and scientific (understanding the relationship between the human visual system and the visual environment). As we learn more about the properties of the human visual system, we will be able to better explain visual behavior. Compared with machine learning-based models, the proposed architecture simulates the low-level features that are common to mammalian cortical architecture, which emerged after millions of years of evolution (i.e., it is not dataset-dependent; Pb, gPb, and the proposed model are compared in Tables 2 and 3). In future research, we will explore the potential role of cells with different functions and integrate them into a unified framework. Different information represents different local information for the input images, and finding an appropriate way to integrate the visual cues will help us understand and simulate the working mode of V1 more deeply. In addition, we will continuously explore more effective and efficient cues, such as luminance contrast in this paper. Such an algorithm might be suitable for some scenarios that require good real-time performance. In summary, the model proposed in this work offers an efficient contour detection solution that uses a simplified DoG function to extract luminance contrast information and a “hard” sigmoid function to obtain an inhibition term. With competitive performance in comparison to the other biologically inspired approaches, we find that the proposed method achieves state-of-art performance for the RuG dataset and near-MCI performance for the BSDSs, with minimal computational complexity.

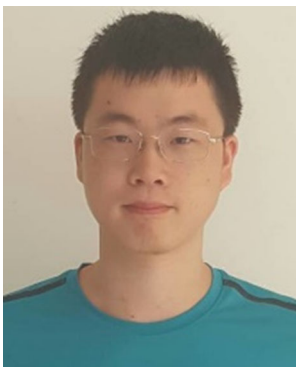
Acknowledgments The authors appreciate the helpful and constructive comments received from the anonymous reviewers of an earlier draft of this paper. This work was supported by the National Natural Science Foundation of China (Grant No. 61866002), Guangxi Natural Science Foundation (Grant No. 2018GXNSFAA138122 and Grant No. 2015GXNSFAA139293), Innovation Project of Guangxi Graduate Education (Grant No. YCSW2018203), and Innovation Project of GuangXi University of Science and Technology Graduate Education (Grant No. GKYC201706 and Grant No. GKYC201803). The funders had no role in the study design; in the collection, analysis, or interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

References

- 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(1):407–430
- Aràndiga F, Cohen A, Donat R, Matei B (2010) Edge detection insensitive to changes of illumination in the image. *Image Vis Comput* 28(4):553–562
- 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
- Azzopardi G, Petkov N (2012) A CORF computational model of a simple cell that relies on LGN input outperforms the Gabor function model[J]. *Biological cybernetics*, 2012, 106(3): 177-189.
- Azzopardi G, Rodriguez-Sanchez A, Piater J, Petkov N (2014) A push-pull CORF model of a simple cell with antiphase inhibition improves SNR and contour detection. *PLoS One* 9(7):e98424
- Canny J (1986) A computational approach to edge detection. *IEEE Trans Pattern Anal Mach Intell* 8(6): 679–698
- 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
- Coen-Cagli R, Dayan P, Schwartz O (2012) Cortical surround interactions and perceptual salience via natural scene statistics. *PLoS Comput Biol* 8(3):e1002405
- Comaniciu D, Meer P (2002) Mean shift: a robust approach toward feature space analysis. *IEEE Trans Pattern Anal Mach Intell* 24(5):603–619
- Cour T, Benezit F, Shi J (2005) Spectral segmentation with multiscale graph decomposition. In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol 2. IEEE, pp 1124–1131
- Dollar P, Tu Z, Belongie S (2006) Supervised learning of edges and object boundaries. In: *CVPR*, vol 2. IEEE, pp 1964–1971
- Fitzpatrick D (2000) Seeing beyond the receptive field in primary visual cortex. *Curr Opin Neurobiol* 10(4): 438–443
- Grigorescu C, Petkov N, Westenberg MA (2003) Contour detection based on nonclassical receptive field inhibition. *IEEE Trans Image Process* 12(7):729–739
- Hubel DH, Wiesel TN (1959) Receptive fields of single neurones in the cat's striate cortex. *J Physiol* 148(3): 574–591
- Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160(1):106–154
- Jones H, Grieve K, Wang W, Sillito A (2001) Surround suppression in primate V1. *J Neurophysiol* 86(4): 2011–2028
- Kapadia MK, Westheimer G, Gilbert CD (2000) Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *J Neurophysiol* 84(4):2048–2062
- Konishi S, Yuille AL, Coughlan JM, Zhu SC (2003) Statistical edge detection: learning and evaluating edge cues. *IEEE Trans Pattern Anal Mach Intell* 25(1):57–74
- Kovesi P (1999) Image features from phase congruency. *Journal of computer vision research. Videre: J. Comp. Vis. Res* 1(3):1–26
- Li C-Y (1996) Integration fields beyond the classical receptive field: organization and functional properties. *Physiology* 11(4):181–186
- Lin C, Xu G, Cao Y, Liang C, Li Y (2016) Improved contour detection model with spatial summation properties based on nonclassical receptive field. *J. Electron. Imaging* 25(4):043018–043018
- Lin C, Xu G, Cao Y (2018) Contour detection model using linear and non-linear modulation based on non-CRF suppression[J]. *IET Image Processing*, 12(6): 993-1003.
- Lin C, Xu G, Cao Y (2018) Contour detection model based on neuron behaviour in primary visual cortex[J]. *IET Computer Vision*, 12(6): 863-872.
- Lindgren JT, Hurri J, Hyvärinen A (2008) Spatial dependencies between local luminance and contrast in natural images. *J Vis* 8(12):6–6
- Liu Y, Cheng M-M, Hu X, Wang K, Bai X (2017) Richer convolutional features for edge detection. In: *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, pp 5872–5881
- Mairal J, Leordeanu M, Bach F, Hebert M, Ponce J (2008) Discriminative sparse image models for class-specific edge detection and image interpretation. In: *European Conference on Computer Vision*. Springer, pp 43–56
- Mante V, Frazor RA, Bonin V, Geisler WS, Carandini M (2005) Independence of luminance and contrast in natural scenes and in the early visual system. *Nat Neurosci* 8(12):1690

28. 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
29. Morrone MC, Owens RA (1987) Feature detection from local energy. *Pattern Recogn Lett* 6(5):303–313
30. Papari G, Petkov N (2011) Edge and line oriented contour detection: state of the art. *Image Vis Comput* 29(2):79–103
31. 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
32. Rand WM (1971) Objective criteria for the evaluation of clustering methods. *J Am Stat Assoc* 66(336):846–850
33. Ren X (2008) Multi-scale improves boundary detection in natural images. In: *ECCV*. Springer, pp 533–545
34. Spratling MW (2013) Image segmentation using a sparse coding model of cortical area V1. *IEEE Trans Image Process* 22(4):1631–1643
35. Tang Q, Sang N, Zhang T (2007) Extraction of salient contours from cluttered scenes. *Pattern Recogn* 40(11):3100–3109
36. Tang Q, Sang N, Liu H (2016) Contrast-dependent surround suppression models for contour detection. *Pattern Recogn* 60:51–61
37. Wei H, Lang B, Zuo Q (2013) Contour detection model with multi-scale integration based on non-classical receptive field. *Neurocomputing* 103:247–262
38. Xiao J, Cai C (2014) Contour detection based on horizontal interactions in primary visual cortex. *Electron Lett* 50(5):359–361
39. 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*. IEEE, pp 2810–2817
40. 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
41. 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
42. Yang K-F, Li C-Y, Li Y-J (2015) Potential roles of the interaction between model V1 neurons with orientation-selective and non-selective surround inhibition in contour detection. *Front. Neural Circuits*, 9, pp. 30
43. Zeng C, Li Y, Li C (2011) Center-surround interaction with adaptive inhibition: a computational model for contour detection. *NeuroImage* 55(1):49–66
44. 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
45. Zhang X-S, Gao S-B, Li R-X, Du X-Y, Li C-Y, Li Y-J (2016) A retinal mechanism inspired color constancy model. *IEEE Trans Image Process* 25(3):1219–1232

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Yijun Cao is pursuing his MS degree in the College of Electric and Information Engineering, Guangxi University of Science and Technology, Liuzhou, China. His area of research is computer vision.



Chuan Lin received his MS degree from the Huazhong University of Science and Technology, China, in 2006. Since 2013, he has been pursuing his PhD at the College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, China. He is a professor in the College of Electric and Information Engineering, Guangxi University of Science and Technology, China. His areas of research are bionic intelligent computing, computer vision, and pattern recognition.



Yi-Jian Pan is pursuing his MS degree in the College of Electric and Information Engineering, Guangxi University of Science and Technology, Liuzhou, China. His area of research is image processing.



Hao-Jun Zhao is pursuing his MS degree in the College of Electric and Information Engineering, Guangxi University of Science and Technology, Liuzhou, China. His area of research is image processing.