

Orientation Histogram-Based Center-Surround Interaction: An Integration Approach for Contour Detection

Rongchang Zhao

Byrons.zhao@gmail.com

*School of Information Science and engineering, Central South University,
Changsha, Hunan 410083, China*

Min Wu

wumin@cug.edu.cn

*School of Automation, China University of Geosciences, Wuhan,
Hubei 430074, China*

Xiyao Liu

xyliu@csu.edu.cn

Beiji Zou

bjzou@csu.edu.cn

Fangfang Li

zhaorc100@163.com

*School of Information Science and engineering, Central South University,
Changsha, Hunan 410083, China*

Contour is a critical feature for image description and object recognition in many computer vision tasks. However, detection of object contour remains a challenging problem because of disturbances from texture edges. This letter proposes a scheme to handle texture edges by implementing contour integration. The proposed scheme integrates structural segments into contours while inhibiting texture edges with the help of the orientation histogram-based center-surround interaction model. In the model, local edges within surroundings exert a modulatory effect on central contour cues based on the co-occurrence statistics of local edges described by the divergence of orientation histograms in the local region. We evaluate the proposed scheme on two well-known challenging boundary detection data sets (RuG and BSDS500). The experiments demonstrate that our scheme achieves a high F -measure of up to 0.74. Results show that our scheme achieves integrating accurate contour while eliminating most of texture edges, a novel approach to long-range feature analysis.

1 Introduction

Contour plays a fundamental and essential role in pattern recognition, especially in object segmentation and recognition of complex scenes. Compared with edge extraction, contour detection is more complicated and high level because it aims to isolate objects accurately and distinguish outlines belonging to different objects while eliminating textures and noncontour edges (Forsyth & Ponce, 2003). Most contour detection algorithms are edge based, which makes use of image gradients to find local luminance changes and more low-level cues.

Numerous efforts relying on luminance changes (Forsyth & Ponce, 2003; Canny 1986; Grigorescu, Petkov, & Westenberg, 2003; Papari, Petkov, & Neri, 2007) have been made to detect contours. Other work (Martin, Fowlkes, & Malik, 2004; Maire, Arbelaez, Fowlkes, & Malik, 2008; Arbelaez, Maire, Fowlkes, & Malik, 2011) has proposed methods to combine multiple local cues into a unified framework for object segmentation and contour detection. Recently, robust contour has been constructed by deep neural networks (Hwang & Liu, 2014; Ganin & Lempitsky, 2014; Kivinen, Williams, & Heess, 2014; Bertasius, Shi, & Torresani, 2015; Shen, Wang, Wang, & Bai, 2015; Xie & Tu, 2015), which model high-level abstractions in images by using multiple processing layers with complex structures.

Contour integration (Field, Hayes, & Hess, 1993; Hess, Hayes, & Field, 2003; Field, Golden, & Hayes, 2014; Vilankar, Golden, Chandler, & Field, 2014; Morimichi & Cornelia, 2016) is an effective approach for boundary detection and object segmentation. This approach holds that contour is a combination of local segments with some consistent constraints; therefore, contour could be generated by grouping and integrating many local low-level features such as contour segments and edges. However, it has been a challenging task for contour integration to conform to a globalization contour architecture based on many local contour segments. Edges are also a key problem.

Fortunately, the human visual system evolved so as to be able to integrate contour segments rapidly and segregate an image into objects or background effectively based on the contour of objects (Albright & Stoner, 2002; Bar, 2004; Landy & Graham, 2004). It has been proposed that neurons in the primary visual system (V1) are not only responsive to visual stimuli and abstract local low-level features, but also group those features from broad parts of the visual field to participate in more complex perceptual tasks such as, object contours perception (Jones, Grieve, Wang, & Sillito, 2001; Kapdia, Westheimer, & Gillert, 2000; Walker, Ohzawa, & Freeman, 2000). For contour detection tasks, the low-level features are extracted and serve as local evidence of the presence of contours and objects. Furthermore, some simple local evidence unites prior knowledge in making use of globalization regularities, for example, the Gestalt principle, and then grouping structural local pieces into a globalization perception.

In this letter, we propose an integration scheme for contour detection based on a novel bio-inspired model, orientation histogram-based center-surround interaction, which mimics the nonlinear long-range interaction among neurons in the primary visual cortex (V1). The basic assumption is that contours can be detected by the globalization integration of structural edges or segments, which are easily detected based on local knowledge, while fractional nonmeaningful texture edges are suppressed by the interaction from neighbors. Furthermore, integration is implemented by means of nonlinear interactions between the center region and its neighbors, and the interaction intensity is held by orientation co-occurrence statistics of neighboring local edges.

First, the scheme takes the raw image as input and creates a multidirectional edges map by computing Gabor convolution energy within local regions. Second, the orientation histogram within a local region centered on each edge point is extracted for the center-surround interactions. Finally, the globalization integration progression is implemented by a nonlinear regression that could accept surrounding interactions to determine the contour intensity of the centered point. In our scheme, the co-occurrence statistics for edges are measured by the divergence of the orientation statistical histogram of local edges.

The rest of this letter is organized as follows. Section 2 examines related work about the bio-inspired contour detection approaches. We introduce our center-surround interaction model in section 3 by presenting our proposed surroundings structure. Section 4 uses the integration approach of contour detection by detailing spatial position weight function, orientation inhibition function, and the co-occurrence statistics of edges. Experimental results and performance evaluation are presented in section 5. Section 6 offers discussion and summary.

2 Related Work

The human visual system can extract object contours effectively, which inspires us to design a biologically motivated algorithm for robust contour detection. Based on the studies of Hubel (1988), and Field et al. (1993), we propose more frameworks to describe the functional architecture for object perception (e.g., Bressloff & Cowan, 2003; Sarti & Citti, 2015; Citti & Sarti, 2006; Sarti, Citti, & Petitot, 2008; Duits & Franklin, 2010a, 2010b).

Physiological research has clearly shown that the area beyond the classical receptive field (CRF) of the majority of neurons in V1, namely, the nonclassical receptive field (non-CRF) or surroundings, although alone is unresponsive to visual stimuli, can exert modulatory effects on the response to the stimuli in the CRF. Usually the modulatory effects that come from the surrounding area of the CRF are called center-surround interaction: the stimuli of the broad area outside the CRF participates in complex nonlinear effects among neuron links that play an important role in representing

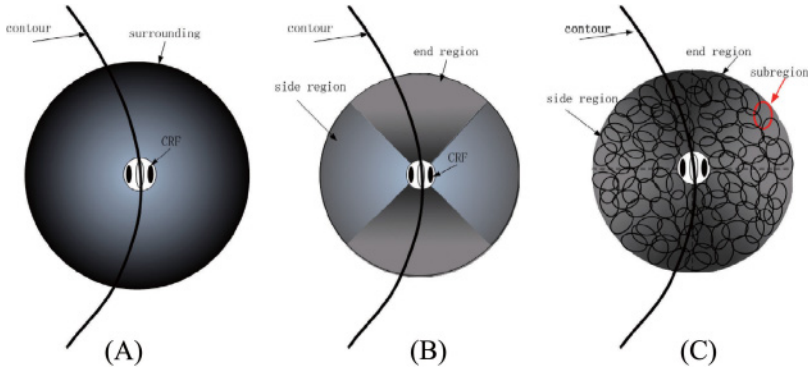


Figure 1: Center-surround interaction model. (A) Classical DOG^+ model. (B) Butterfly-shaped model. (C) Our new model where the surrounding comprises many independent subregions.

visual information and featuring integration in a broad range (Jones et al., 2001; Kapadia et al., 2000; Walker et al., 2000; Hubel & Wiesel, 1968; Li & Li, 1994; Li, 1996).

In a contour extraction task, the modulatory effect makes use of the co-occurrence statistics between the central area and its surroundings to suppress noncontour edges while keeping the integral contour. In general, the inhibition exerts effects in a ring-formed surrounding area. Furthermore, physiological findings show that the intensity of the inhibition decreases as the distance from the concerned point of receptive field increases. Therefore, this form of inhibition can be realized using the difference of gaussian function, known as the difference of gaussian (DOG) model shown in Figure 1A. However, the DOG model, which cannot discriminate an object contour from noncontour edges, is invariable and insensitive to the input image pattern, especially the gradient direction in the local region. Subsequently, electrophysiological studies (Li & Li, 1994; Li, 1996) have shown that the non-CRF area could be described as four symmetrical parts according to the stimuli orientation at the center point: two end regions and two side regions shown in Figure 1B. It is called a butterfly-shaped model and is flexible in analyzing and understanding different patterns of input images.

Based on the idea of center-surround interaction, several inhibition-based contour detection models have been proposed (Grigorescu et al., 2003; Papari et al., 2007; Ursino & La Cara, 2004; La Cara & Ursino, 2008; Long & Li, 2008; Petkov & Westenberg, 2003; Ren, 2008; Tang, Sang, & Zhang, 2007). The typical model, proposed by Grigorescu et al. (2003) and Petkov and Westenberg (2003) and named the isotropic and anisotropic inhibition models, employs a ring-formed area for surrounding structures and the DOG model for computing the inhibition effects. Their models enhance

contour detection by eliminating a few local edges of textured background and make it more effective than the classical Canny edge detector (Canny, 1986). However, the fact that the consistent inhibition affects contour and texture uniformly leads to a serious issue where information in the surrounding areas always exerts equivalent suppression on edges belonging to object contours. This issue, known as self-inhibition, not only suppresses texture but also makes weak contours difficult to pop out. To resolve the issue of self-inhibition, researchers have to consider the balance of inhibitions between object contours and background textures. Grigorescu et al. (2007) have changed the surrounding structure by splitting the inhibition into two truncated half-rings, leaving a band region oriented along the concerned edge with no inhibition. In addition, Papari and Petkov (2011) developed an operator based on steerable filters for computing inhibitions to overcome self-inhibition.

Based on the butterfly-shaped surroundings hypothesis, numerous inhibition-based contour detection algorithms (Long & Li, 2008; Tang et al., 2007; Zeng, Li, & Li, 2011; Zeng, Li, Yang, & Li, 2011) have been proposed. Those methods adopt different inhibition in the end and side regions of surroundings to restrain background textures while retaining edges belonging to the contours of objects. Furthermore, several approaches (La Cara & Ursino, 2008; Ren, 2008; Wei, Lang, & Zuo, 2013) exert inhibition in different scales to realize more effective contour detection through integrating multiscale, multiregion information. Yang, Li, and Li (2014) proposed a multifeature-based center-surround framework in which multiple features are combined according to a scale-guided strategy, to improve the performance of perceptually salient contour detection. Aiming to combine brightness and color features to maximize the reliability of contour detection, a framework based on color-opponent mechanisms was proposed (Yang, Gao, Li, & Li, 2013; Yang, Gao, Guo, Li, & Li, 2015).

Although the accurate physiological structure of non-CRF is unknown, there is no doubt that different types of non-CRF exist, and they are flexible enough to analyze and understand visual information. Therefore, it is clear that the adaptive interaction among different areas of receptive fields of V1 neurons is one of the foundations for V1 to capture and group different patterns of visual stimuli effectively.

The model we propose in this letter is composed of a central region and surrounding subregions. Each subregion extracts a local image pattern and exerts a flexible nonlinear modulatory effect on the response of the central region, with the modulation intensity determined by orientation co-occurrence statistics in the local region. Mathematically, the center-surround interaction is implemented by a nonlinear regression learned from the statistical divergence of an orientation histogram of local contour segments. Depending on the interactions between each central region and its surrounding subregions, the model groups structure local segments into contours while suppressing texture edges.

3 The Center-Surround Model

Our model evolves from both the ring-shaped and butterfly-shaped surroundings, which are not useful for contour integration. They cannot keep a consistent inhibition in a broad area without distinguishing the contextual configuration flexibly even though they have achieved some effective results for contour detection.

The hypothesis behind our model, illustrated in Figure 1C, is that the surroundings of each central region have a number of independent subregions that can extract local low-level image features like edges, gradient orientation, energy, and contrast. These features represent the presence of contour segments in the local region. Moreover, the subregion exerts a nonlinear modulatory effect on the response of the central local region. The nonlinear modulation is used to identify consistent local contour segments and align them to provide powerful evidence of the contour presence since the surroundings structure composed of subregions could implement flexible interaction among neighboring neurons.

3.1 Local Contour Segments. The response of the simple neuron in the visual system represents local patterns that act as fundamental elements of visual objects. The local low-level features like edges, gradient orientation, energy, and contrast provide evidence of the object contour. In the hierarchical architecture, the low-level features are grouped into high-level object blocks for visual recognition. Consequently, the primary function of our model is to extract the contour segments based on local low-level image features.

The contour is always caused by abrupt changes and discontinuities of the input luminance profile, which can be detected as points of high gradient magnitude. To simplify our model, Gabor filters are used to extract the local segments, and Gabor energy is employed to represent the intensity of local segments in the central region. A two-dimensional Gabor filter can be expressed as

$$g_{\sigma, \lambda, \varphi, \theta}(x, y) = \exp\left(-\frac{\tilde{x}^2 + \gamma^2 \tilde{y}^2}{2\sigma^2}\right) \cos\left(2\pi \frac{\tilde{x}}{\lambda} + \varphi\right), \quad (3.1)$$

$$\tilde{x} = x \cos \theta + y \sin \theta, \quad \tilde{y} = -x \sin \theta + y \cos \theta, \quad (3.2)$$

where θ is the preferred orientation of the CRF and $\gamma = 0.5$ is a constant since it is the reasonable value for the receptive field of the simple cells in cat striate cortex, called the spatial aspect ratio, that determines the ellipticity of the receptive field. $g(x, y)$, with $\varphi = 0$ and $\pi/2$ representing the even and odd Gabor filters respectively. σ , the standard deviation of the gaussian factor, defines the size of the Gabor-simulated CRF orthogonal to the preferred orientation of CRF. λ is the wavelength and often represents

the spatial frequency bandwidth as σ/λ . In our model, σ/λ is set to the value 0.56 based on the biological findings. Note that we choose σ as a free parameter since σ/λ and λ are determinate.

Response $E(x, y)$ to input $f(x, y)$ with luminance distribution is given by convolution,

$$e_{\sigma,\gamma,\varphi,\theta}(x, y) = f(x, y) \otimes g_{\sigma,\gamma,\varphi,\theta}(x, y), \quad (3.3)$$

$$E_{\sigma,\gamma,\theta}(x, y) = \sqrt{e_{\sigma,\gamma,0,\theta}^2(x, y) + e_{\sigma,\gamma,-\pi/2,\theta}^2(x, y)}, \quad (3.4)$$

where \otimes denotes the convolution operator to the effect of the Gabor filter on the input $f(x, y)$. Hence, $e_{\sigma,\lambda,0,\theta}(x, y)$ and $e_{\sigma,\lambda,-\pi/2,\theta}(x, y)$ are the responses of even and odd Gabor-stimulated neurons, respectively. The Gabor energy distributes in a number of orientations because the Gabor filters are orientation sensitive. Then the preferred orientation map $\tilde{\theta}$ and the maximum Gabor energy \tilde{E} are given by

$$\theta_i = \frac{(i-1)\pi}{N_\theta}, \quad i = 1, 2, \dots, N_\theta, \quad (3.5)$$

$$\begin{aligned} \tilde{\theta}_{\sigma,\lambda}(x, y) &= \arg \max \{E_{\sigma,\lambda,\theta_i}(x, y) | i = 1, 2, \dots, N_\theta\}, \\ \tilde{E}_{\sigma,\lambda}(x, y) &= \max \{E_{\sigma,\lambda,\theta_i}(x, y) | i = 1, 2, \dots, N_\theta\}. \end{aligned} \quad (3.6)$$

Here, N_θ is the number of orientation. To balance the running efficiency and effect, in this work we choose 45 different orientations with equal intervals, so N_θ is set to be 45. The local low-level contour cues are represented by the preferred orientation map and corresponding Gabor energy map.

3.2 Surroundings Structure. To achieve flexible and self-adaptive surrounding interaction, our hypothesis about the surroundings of a central region is that the surrounding area is composed of several subregions that have independent perceived capability of local low-level features, and inputs of those subregions exert different effects on the output of the central local region, respectively. It is reasonable to assume that the broad receptive field could flexibly perceive more complete visual appearances based on variable modulations and self-adaptive interaction between the central and surrounding regions.

First, we define the surroundings as a series of feature-receptive units to capture local low-level features in our model. We refer to this model as a subregion. Although the structure of subregions is similar to the central region, the former perceives features of the local inputs and exerts a non-linear interaction on the local response of its central region. To capture the orientation information, the subregion of the surrounding area is designed

as

$$\mathfrak{N} : \{(\tilde{x} - \tilde{x}_0)^2 + \gamma^2(\tilde{y} - \tilde{y}_0)^2 \leq \sigma_2^2 | \tilde{x}, \tilde{y} \in \mathbb{Z}\}, \quad (3.7)$$

$$\text{where } \tilde{x}_0^2 + \tilde{y}_0^2 > \sigma_1^2 \quad (3.8)$$

$$\tilde{x}(\tilde{x}_0) = x(x_0) \cos \theta + y(y_0) \sin \theta, \quad \tilde{y}(\tilde{y}_0) = -x(x_0) \sin \theta + y(y_0) \cos \theta. \quad (3.9)$$

The subregion is an ellipse-shaped area, and the direction θ of the ellipse is the same as the preferred orientation of the center of the subregion: γ is the spatial aspect ratio that determines the eccentricity of the ellipse, and σ_1 determines the size of the ellipse-shaped subregion, while σ_2 determines the size of the circle-shaped central region. $(\tilde{x}_0, \tilde{y}_0)$ is the coordinate of the center of an ellipse-shaped subregion \mathfrak{N} .

Second, unlike other approaches, the subregion in our model possesses the perception of local low-level contour cues independently. The perception is designed as Gabor-simulated form, which is the same as the function of the central region since the center-surround interaction is based on a comparison of local feature between subregions and its central region. The perception of the ellipse-shaped subregion \mathfrak{N} is given by equations 3.5 and 3.6. In the same way as the capture process in the central region, the subregion perceives local contour cues using the convolution of stimuli in the ellipse-shaped subregion with Gabor functions and captures the preferred orientation and the corresponding Gabor energy map in the subregion.

Based on the value of σ_1 in the subregion model, the surroundings are divided into several ellipse-shaped subregions (see Figure 1C), which can capture the local image pattern of the inputs. The size of the subregion is adaptive to the pattern of the input image. To simplify the model, we choose a fixed value for the size of subregions in the surroundings for this letter.

The surroundings \mathbb{S} is expressed as

$$\mathbb{S} = \{\mathfrak{N}_i | \cap \mathfrak{N}_i \neq \emptyset, \cup \mathfrak{N}_i = \mathbb{S}, i = 1, 2, \dots\}. \quad (3.10)$$

This means that all the subregions can cover all the points of the surroundings, and it does not matter that subregions may be overlapping each other.

4 Orientation Histogram-Based Interaction and Contour Detection —

According to the findings of physiological research, the types of neurons with inhibitory effects at both the end and side regions show weak or no response to broad field homogeneous textures but respond vigorously to various textural contrasts (Zeng, Li, & Li, 2011). Consequently, this feature-selective inhibitory effect was viewed as the neural basis for our contour detection model.

In the contour detection algorithm, the contour intensity in the central region is determined by the local contour cues and interactions with the surrounding subregions depending on the segment's co-occurrence statistics. In addition, the interaction is weighted by a spatial position function and an orientation inhibition function. In short, the interaction is determined based on spatial position, orientation inhibition, and the co-occurrence statistics. (We discuss each part of the contour detection algorithm in section 4.1.)

We simulate the center-surround interaction for contour detection using the model given as

$$r(x, y) = H\left(\tilde{E}_{\sigma, \lambda}(x, y) - \sum_i \alpha_i M_i(x, y; \sigma, \lambda)\right), \quad (4.1)$$

where $r(x, y)$ is the final output of the contour detector proposed in our work at location (x, y) , and $\tilde{E}_{\sigma, \lambda}(x, y)$ is the response to the input raw image presented in the local region given by equation 3.6. $M_i(x, y; \sigma, \lambda)$ denotes the modulatory term from the i th subregion of the center region, which is determined by spatial position function, orientation inhibition function, and the co-occurrence statistics introduced in section 4.1. The factor α is the modulation coefficient that controls the linking intensity of the modulation from the subregion. $H(s) = \max(s, 0)$ is a half-wave rectification operator simulating that the cortical cell can fire only when positive excitation is received. Consequently, the final output $r(x, y)$ is represented by a rectified linear regression of the local response and its surrounding interactions. Equation 4.1 indicates that the contour intensity presented in the center region is modulated by the nonlinear combination of several subregions within its surroundings. As a result, the center-surround interaction is transformed as a rectified linear regression; therefore, the parametric α can be learned by the gradient descent algorithm (Boyd & Vandenberghe, 2004).

4.1 Spatial Position Weight Function. Physiological findings indicate that the modulatory effect from the surroundings decreases with increasing distance from the center of the receptive field (Jones et al., 2001; Kapadia et al., 2000; Walker et al., 2000; Li, 1996). Much like other approaches, we adopt the location weight model established by Grigorescu et al. (2003), but the spatial position weight is applied only to the surroundings. For a given point (x, y) in an image, the spatial position weight function $w_D(x, y)$ is given by

$$w_D(x, y; \sigma, k) = \frac{DOG^+(x, y; \sigma, k)}{\|DOG^+(x, y; \sigma, k)\|_1}, \quad (4.2)$$

$$\begin{aligned} &DOG^+(x, y; \sigma, k) \\ &= H\left(\frac{1}{2\pi(k\sigma)^2} \exp\left(-\frac{x^2 + y^2}{2(k\sigma)^2}\right) - \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)\right), \end{aligned} \quad (4.3)$$

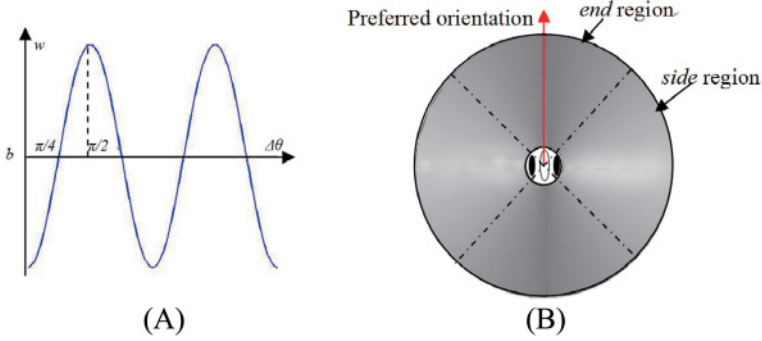


Figure 2: (A) Curve of weight function. (B) Schematic plan of the weight function.

where $\|\cdot\|_1$ denotes the L_1 norm and $H(s) = \max(s, 0)$. Thus, DOG^+ defines a ring-shaped region and is formulated by the difference of two gaussian functions, so the location weight model is called difference-of-gaussian (DOG) model. k is the ratio of the standard divisions of two gaussian functions; in most research, k is set as 4 based on the neurophysiological findings.

4.2 Orientation Inhibition Function. Unlike the existing model, the orientation inhibition function in our model is used to simulate the modulatory effect from the side and end regions. Here, the modulatory effect related to contour orientation is represented as a continuous curve given as

$$w_\theta(x, y) = \frac{1}{C}[\cos(\pi + 2\Delta\theta(x, y)) + b], \quad (4.4)$$

where $\Delta\theta(x, y)$ is the orientation difference between the (x, y) point orientation of surroundings and the preferred one of the center and b defines the bias of the weight. The numerator C is used to normalize the weight. It is the end region when $w_\theta(x, y) > b/C$, and the side region elsewhere.

The curve of the inhibition function (see Figure 2A) illustrates that the weight of the modulatory effect increases with $\Delta\theta$ is from the minimum when $\Delta\theta$ is equal to 0 to the maximum while $\Delta\theta$ is valued as $\pi/2$. To simplify the interpretation of the model, we choose the parameter b as C . The extreme values of inhibition function are located at the integer multiples of $\pi/2$ (e.g., 0, $\pi/2$, and π), and it possesses positive value when $\Delta\theta$ is on the intervals $[\pi/4, 3\pi/4]$ and $[5\pi/4, 7\pi/4]$. This means that the region of surroundings corresponding to the two intervals exerts a more powerful inhibition on the response to input at the center than other regions when the model is considered an inhibition one. The region corresponds to the side region in the butterfly-shaped model. By the crossing point of the weight

function with b , the surroundings can be divided into four regions. It is the end regions when $\Delta\theta$ is on the intervals $[-\pi/4, \pi/4]$ and $[3\pi/4, 5\pi/4]$, where the area exerts weak inhibition because edges belonging to object contours generally traverse the center and the end regions of surroundings simultaneously. The side regions are defined by $\Delta\theta$ on the intervals $[\pi/4, 3\pi/4]$ and $[5\pi/4, 7\pi/4]$. However, the weight function in our model, which is different from the butterfly-shaped model, varies with the location of each point of the surroundings; furthermore, the transition from one region to another is smooth. In this study, we call this weight function an orientation inhibition function.

Together with the spatial position weight function $w_D(x, y)$ and the orientation inhibition function $w_\theta(x, y)$, the total weighting function is given by

$$\begin{aligned} w(x, y) &= w_D(x, y) \cdot w_\theta(x, y) \\ &= \frac{DOG^+(x, y; \sigma, k)}{\|DOG^+(x, y; \sigma, k)\|_1} \cdot \frac{1}{C} [\cos(\pi + 2\Delta\theta(x, y)) + b], \end{aligned} \quad (4.5)$$

where DOG^+ is described by equation 4.3.

4.3 Orientation Histogram-Based Interaction. The key issue of contour integration is aligning local contour segments to form the globalization contour architecture based on the co-occurrence statistics of contour segments. Inspired by the HOG algorithm (Dalal & Triggs, 2005), which counts occurrences of gradient orientation in the local position of an image, we assume that the distribution of the preferred orientation of segments in a local region can describe the local appearance and characteristics of contours. Consequently, the statistical divergence of the distribution can also represent the co-occurrence of contour segments, which implies the structural consistency of the contours. Thus, we first consider the distribution of preferred orientation; then we calculate the statistical divergence of the distribution to control the interaction of the local regions.

4.3.1 Orientation Histogram. The orientation histogram is calculated based on the directional Gabor energy within the surroundings of the central region and relies on the preferred orientation and corresponding Gabor energy. The first step is to create the histogram of the preferred orientation in each subregion. For the subregion within surroundings, a local one-dimensional histogram of the preferred orientation and Gabor energy is accumulated over points of the subregion. Each point of the subregion calculates a weighted vote for an orientation-based histogram channel based on the preferred orientation of the point in the subregion, while the vote is a function of the Gabor energy itself. The votes are accumulated in the orientation bin over the subregion within the CRF surroundings. There are

nine, histogram channels, and the orientations from 0 to π are divided into nine channels, so each channel possesses 20 degrees in our work because the unsigned channel performs best experimentally. To compute the orientation binning, the role of the subregion and surroundings of our model is the same as the cell and block in the HOG algorithm.

To reduce aliasing, votes are interpolated bilinearly between the neighboring bin centers in orientation. Let $center_b_t$ denote the center of the bin that the preferred orientation falls in. Therefore, votes belonging to this bin are calculated as

$$h(t, x, y) = Gabor_energy(x, y) \cdot |orientation(x, y) - center_b_t| / bin_width, \quad (4.6)$$

and votes on the neighboring bin are calculated as

$$h(s, x, y) = Gabor_energy(x, y) \cdot |orientation(x, y) - center_b_s| / bin_width \quad (4.7)$$

$$s = \begin{cases} t - 1, & \text{if } orientation(x, y) < center_b_t \\ t + 1, & \text{if } orientation(x, y) \geq center_b_t \end{cases}, \quad (4.8)$$

where $Gabor_energy(x, y)$ and $orientation(x, y)$ are the Gabor energy and the preferred orientation of the point (x, y) in the subregion, respectively; t and s are the index of the neighboring bins; and bin_width denotes the width of each bin. The histogram of the preferred orientation is represented by the histogram h for each subregion of the surroundings.

4.3.2 Interactions Based on Co-Occurrence Statistics. The interaction between subregions and its central region is controlled by the co-occurrence statistics of edges, represented with the help of the orientation histogram divergence between local regions. From the point of view of information entropy, the orientation histogram of the subregion provides information gain to discriminate contour intensity. Therefore, in this work, the statistical divergence of orientation histogram is measured by the Kullback-Leibler divergence function, given as

$$D_{KL}(h(x_c, y_c), h(x, y)) = \sum_i \ln \left(\frac{h(i, x_c, y_c)}{h(i, x, y)} \right) h(i, x, y), \quad (4.9)$$

where $D_{KL}(h(x_c, y_c), h(x, y))$ is the Kullback-Leibler divergence function of $h(x, y)$ from $h(x_c, y_c)$, which often is used to measure the difference between two distributions in information theory. $h(x, y)$ and $h(x_c, y_c)$ are the histograms of the preferred orientation of local edges in the subregions of

the surroundings and the center region, respectively, where (x_c, y_c) denotes the central coordinate of the center region.

The modulatory term $M_i(x, y; \sigma, \lambda)$ in equation 4.1 indicates the intensity of interaction between the i th subregion and its central region. The modulatory term is determined by spatial position function, orientation inhibition function, and the co-occurrence statistics. Here, we consider the interaction as a kernel function $k_\sigma(x, y)$:

$$k_\sigma(x, y) = w(x, y) \cdot D_{KL}(h(x_c, y_c), h(x, y)). \quad (4.10)$$

Consequently, the modulatory term $M_i(x, y; \sigma, \lambda)$ can be computed by a convolution of the Gabor energy function $\tilde{E}_{i,\sigma,\lambda}(x, y)$ with the kernel function $k_\sigma(x, y)$ as

$$M_i(x, y; \sigma, \lambda) = \tilde{E}_{i,\sigma,\lambda}(x, y) \otimes k_\sigma(x, y). \quad (4.11)$$

4.4 Contour Detection Algorithm. Based on the center-surround interaction model, we propose a novel contour detection algorithm in this section. Figure 3 shows the schematic drawings illustrating the general flowchart of the proposed algorithm for contour detection. Given a raw image, Gabor energy maps and the preferred orientations are calculated by convolution with a series of directional Gabor functions. The energy maps indicate the appearance of local edges and textures at different orientations, spatial frequencies, and phases. For each pixel of the image, the central region and surrounding subregion are defined. Then the histograms of the preferred orientation are counted in each central region and subregion. Based on the divergence of the orientation histogram, adaptive modulations are calculated for each subregion. The modulatory effects from every subregion are exerted on the Gabor energy $\tilde{E}_{\sigma,\lambda}(x, y)$ with a weight function $w(x, y)$. The modulation is a subtraction operation when the surrounding is inhibitive and an addition operation when facilitative. Finally, to construct the binary contour map, the well-known nonmaximum suppression is used. For the contour integration and detection algorithm, see algorithm 1.

5 Performance Evaluations

In this section, we present the results of a number of experiments. To evaluate the robustness of the proposed contour detection method, the model with various combinations of specific parameters is tested on the database RuG data set (Grigorescu et al., 2003) and the performance is shown with box-and-whisker diagrams. Then we give benchmarking results of the contour detection algorithm compared with some other bio-inspired or other

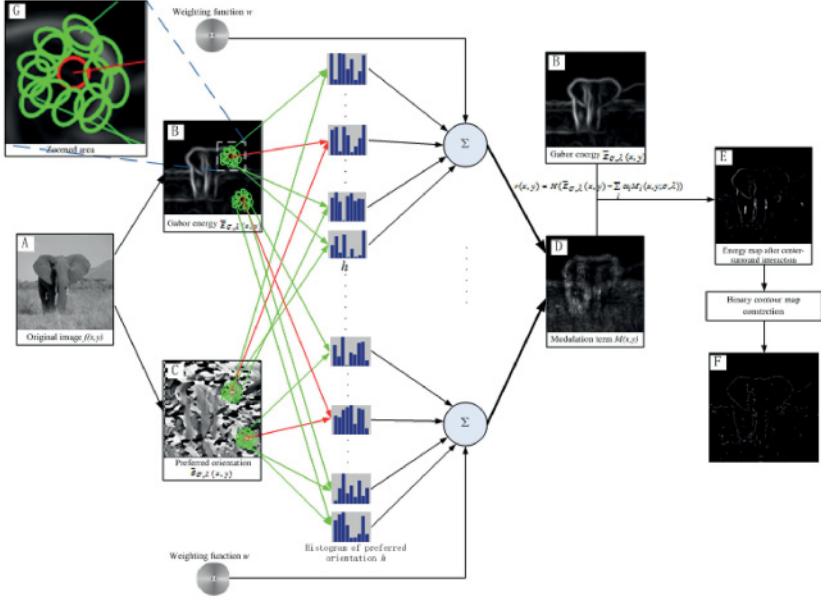


Figure 3: Schematic plan of our algorithm. (A) Original input image $f(x, y)$. (B) The maximum Gabor energy map computed by the convolution of $f(x, y)$ with Gabor functions. (C) The corresponding preferred orientation map. (D) Modulation term $M_i(x, y)$ from the i th subregion of surroundings. (E) The energy map after modulations from surrounding subregions. (F) The binary contour map constructed by a process of binarization based on the modulated energy map. (G) The zoomed area of B.

methods on the challenging database BSDS500 (Martin, Fowlkes, Tal, & Malik, 2001).

5.1 Robust Evaluation. We adopt the measure method proposed by Grigorescu et al. (2003) on the RuG data set of images that has been used to evaluate contour detector based on the ground-truth contour maps drawn manually. For the ground-truth contour maps, E_{GT} denotes the set of contour pixels of the ground-truth map, E denotes the set of correctly detected contour pixels, E_{FP} denotes the set of pixels detected by a detector but not belonging to ground truth, and E_{FN} denotes the set of pixels of ground-truth contour that the detector missed. The performance measure P is computed as

$$P = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FP}) + \text{card}(E_{FN})}, \quad (5.1)$$

where $\text{card}(S)$ denotes the number of the set S .

Algorithm 1: Contour Detection.

Calculate Gabor energy maps and the preferred orientations for each pixel with Gabor functions $g_{\sigma,\lambda,\phi,\theta}(x, y)$

For each pixel (x, y)

 Define subregions for the surroundings

 Compute weighting function $w(x, y)$

 Count its histogram of preferred orientation h_C for its center region

For each subregion indexed i

 Count the histogram of preferred orientation h_{sub_i}

 Calculate the DL divergence D_{KL} of h_{sub_i} from h_C

 Calculate the modulation term $M_i(x, y)$ of subregion i and

 subtract/add it from/to the Gabor energy $g_{\sigma,\lambda,\phi,\theta}(x, y)$ of pixel (x, y)

End

End

The nonmaxima suppression followed by hysteresis thresholding

Performance is usually affected by the parameters of the model. To evaluate robustness, we test the model with various combinations of parameters. For Gabor functions in our model, there is only a single parameter to be tuned, the deviation σ of Gabor filters in equation 3.1. We use $\sigma \in \{1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4\}$. To simplify the test, for every subregion, we adopt the fixed value for the modulation coefficient $\alpha \in \{1.0, 1.2\}$ in equation 4.1. For the last part of our contour detection algorithm where nonmaxima suppression is used for binary contour map construction, we use $p \in \{0.5, 0.4, 0.3, 0.2, 0.1\}$.

For each of the 40 test images in the RuG data set, the 80 P values computed with the 80 groups of parameter combinations of σ , α , p are statistically illustrated in Figure 4 by using statistical box-and-whisker diagrams. Each whisker represents the range of the P value with 80 groups of parameter combinations for each image, the top end of each whisker represents the optimal performance, and the horizontal red line shows the median value of the performance. In the graph, the results of our model are compared with most of the surround inhibition-based contour approaches

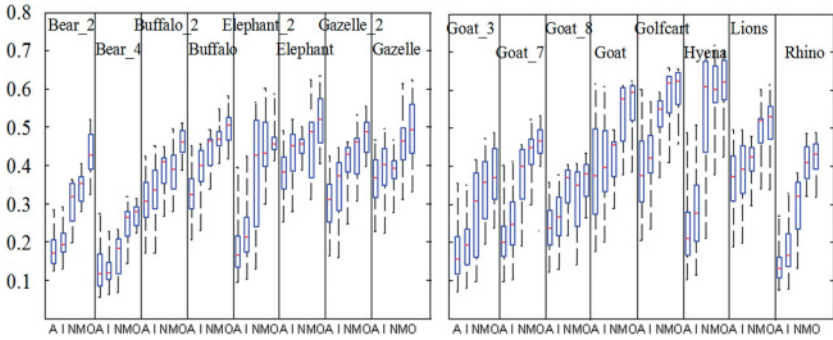


Figure 4: Box-and-whisker plots of the contour detection performance of the anisotropic inhibition model (denoted by A), the isotropic inhibition model (denoted by I; Grigorescu et al., 2003, and Petkov & Westenberg, 2003), the adaptive inhibition model (denoted by N; Zeng, Li, & Li, 2011; Zeng, Li, Yang et al., 2011), the multiscale integration model (denoted by M; Wei et al., 2013) and our model (denoted by O) for the 40 images of the RuG data set used in this letter. The box-and-whisker results of M are from Wei et al. (2013); we implemented the others, so the results presented for M may be not exactly identical to the original paper. However, that difference does not affect the performance comparison.

proposed recently, for example, the standard anisotropic and isotropic inhibition models (Grigorescu et al., 2003; Petkov & Westenberg, 2003), the butterfly-shaped inhibition model (Zeng, Li, & Li, 2011; Zeng, Li, Yang et al. (2011), and the multiscale integration model (Wei et al., 2013). From the plot, we can see that our model achieves better performance than the other methods. For most of test images, the whisker's top ends of our model are obviously higher than those of the other models, which reveal that our model could obtain the optimal performance for contour detection compared with the other methods. In addition, the whiskers of most images are compact because our model is insensitive to parameters. Finally, our model produces a better median value for most cases.

In Figure 5, we compare our model with most of the bio-inspired models. Those contour maps are the best in the results tested with many groups of parameter combinations. The rows from top to bottom show the original images, corresponding ground truth, results from anisotropic and isotropic inhibition models (Grigorescu et al., 2003; Petkov & Westenberg, 2003), the adaptive inhibition model (Zeng, Li, & Li, 2011; Zeng, Li, Yang et al., 2011), the multiscale integration model (Wei et al., 2013), and our model, respectively. The results clearly show that our model could effectively suppress noncontour edges and background textures and perform better than the other models in detecting object contour from cluttered background.

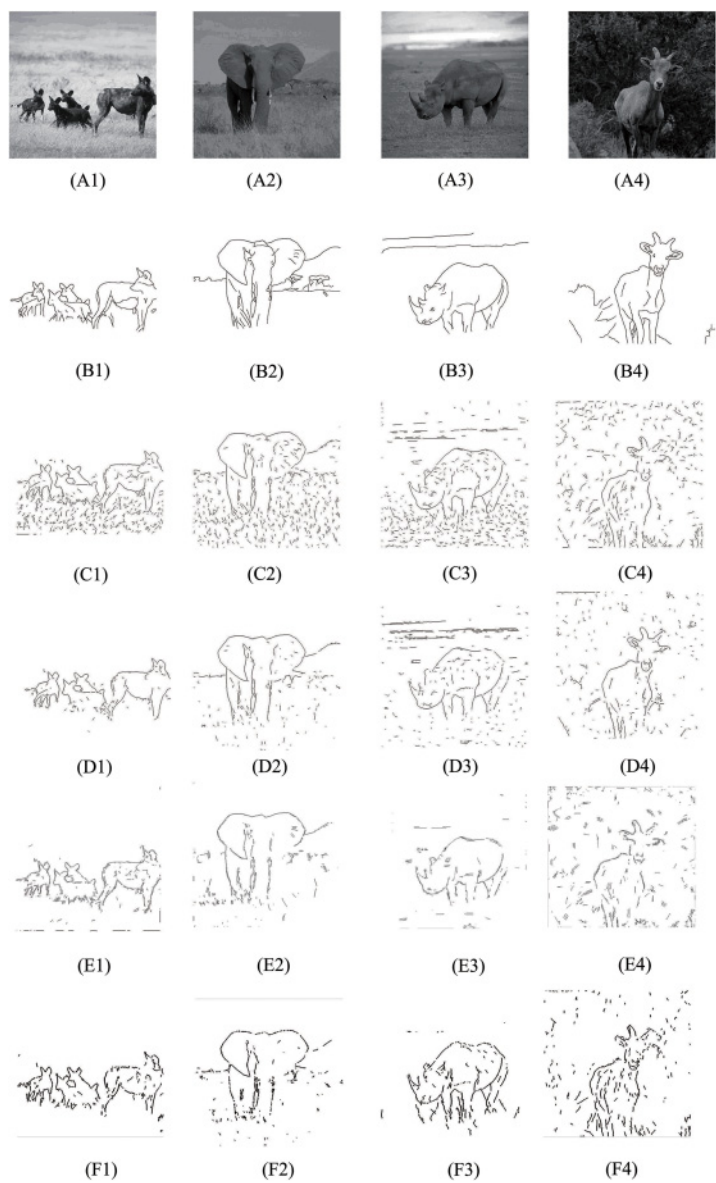


Figure 5: Images from the RuG data set and corresponding contour map obtained by different algorithms. The rows from top to bottom represent (A) the original image, (B) the ground truth, (C) the contour map from the anisotropic inhibition model (Grigorescu et al., 2003), (D) the adaptive inhibition model (Zeng, Li, & Li, 2011), (E) the multiscale integration model (Wei et al., 2013), and (F) our model.

Furthermore, the contour maps of our model retain as many as possible integral contours of tiny objects, for example, the eyes of those animals and the bird in the second image, which are lost in other methods.

5.2 Benchmarking Results. Contour detection performance is evaluated in the precision-recall framework (Jesse & Mark, 2006). The precision-recall curve captures the trade-off between accuracy and noise. High precision means that an algorithm returns more true positives than false ones, and high recall means that an algorithm returns more true positives than false negatives. Precision and recall are computed based on the result maps of the model and ground truths. For contour detection, high precision indicates that the detected contour map contains more contour pixels belonging to ground truth than noise, while high recall suggests that the map contains most of contour pixels belonging to ground truth. In the performance evaluations of contour detection, high precision and recall are expected simultaneously, but more noise would be introduced when high recall is obtained. Generally, a convex precision recall curve is expected. In addition, the F -measure combines precision and recall as $F = PR/(\alpha R + (1 - \alpha)P)$. Here, the larger the F , the better.

In Figure 6, we compare some of previous contour detectors. We choose state-of-the-art methods to compare more extensively—for example Crisp Boundaries (Isola, Zoran, Krishnan, & Adelson, 2014), gPb-owt-ucm detector (Arbelaez et al., 2011), Sketch Tokens (Lim, Zitnick, & Dollar, 2013), Normalized Cut (Cour, Benezit, & Shi, 2005), Canny algorithm (Canny, 1986), and other contour detection methods (Dollar & Zitnick, 2013; Ren & Bo, 2012; Comaniciu & Meer, 2002; Felzenszwalb & Huttenlocher, 2004). The results of the comparisons are based on the widely used benchmark for the contour detection algorithm, BSDS500, the Berkeley segmentation data set, which contains 500 test images and corresponding ground truth. We report performance by the precision-recall curve and F -measure, which show that our approach to contour detection can match the state-of-the-art.

6 Discussion and Conclusion

Most bio-inspired contour detection models are based on center-surround interaction of the V1 neurons, but those models are so inflexible that they cannot accurately distinguish contours from edges and textures. This letter proposes a novel scheme to group structural edges into contours by the orientation-based center-surround interaction model, which integrates structural edges or contour segments into a globalization contour architecture based on edge co-occurrence statistics. Noncontour edges and textures are suppressed effectively through a nonlinear regression from surroundings. The experimental results exhibit its superior performance in contour detection.

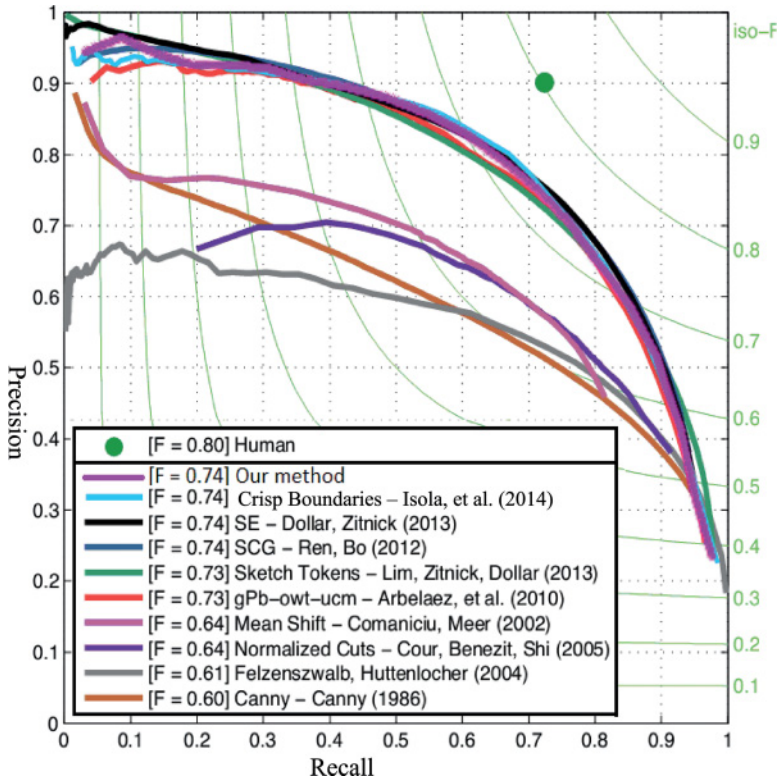


Figure 6: The precision-recall curve of contour detection on the BSDS 500 Benchmark. The isolated green dot refers to the human performance assessed on the same image. Source: Isola et al. (2014) with our results added.

The scheme is robust to suppress noncontour edges and textures for three reasons. First, our proposed center-surround interaction model has a flexible structure. There is no strict definition of the side and end regions. Rather, side and end are represented by the cosine-simulated weight function in equation 4.4 and Figure 2, which changes smoothly in each region and over the regions. This means that the modulation intensity is multi-valued rather than fixed states presenting inhibition or not. Also, the surroundings of a V1 neuron's receptive field are made up of many subregions that could perceive local low-level cues of contours independently, such as edges and contour segments. In other words, the proposed model can achieve pixel-wise contour detection.

Second, the long-range interaction is described as a nonlinear modulation from the surrounding area on the response of the central region. The modulatory effect is determined by the partial position, orientation

inhibition, and co-occurrence statistics of the edges. Consequently, the contour detector perceives information from a large range of vision and groups the aligned edges into structural contours.

Finally, although some content of our model is speculative, there is physiological evidence supporting our basic idea about center-surround interaction. Increasing studies show that the V1 neuron acts as an adaptive information processor, and this kind of processing is dependent on visual content. Furthermore, the adaptive performance originates from the dynamic center-surround interaction (Barthelemy, Vanzetta, & Masson, 2006).

Our task is to develop an integration approach to contour detection. Although the proposed approach is useful for contour detection, it is extremely difficult to achieve such accurate judgment about the noncontour edges and how to deal with the jag of contours. More work could be done, for example, on the prediction of continuous contours and contour integration on a different scale. This will be the subject for our further studies.

Acknowledgments

This work was supported by the scientific research foundation of Central South University. We thank N. Petkov et al. and J. Malik et al. for their RuG data set and BSDS500 benchmarks. We are very grateful to the anonymous reviewers for their valuable comments and suggestions on the manuscript and to Kun Zhan, Yide Ma, and Hongjuan Zhang for many useful suggestions.

References

- Albright, T. D., & Stoner, G. R. (2002). Contextual influences on visual processing. *Annual Review of Neuroscience*, 25, 339–379.
- Arbelaez, P., Maire, M., Fowlkes, C., & Malik, J. (2011). Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5), 898–916.
- Bar, M. (2004). Visual objects in context. *Nature Reviews: Neuroscience*, 5(8), 617–629.
- Barthelemy, F. V., Vanzetta, I., & Masson, G. S. (2006). Behavioral receptive field for ocular following in humans: Dynamics of spatial summation and center-surround interactions. *Journal of Neurophysiology*, 95(6), 3712–3726.
- Bertasius, G., Shi, J. B., & Torresani, L. (2015). DeepEdge: A multi-scale bifurcated deep network for top-down contour detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4380–4389). Piscataway, NJ: IEEE.
- Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge: Cambridge University Press.
- Bressloff, P. C., & Cowan, J. D. (2003). The functional geometry of local and long-range connections in a model of V1. *Journal of Physiology Paris*, 97(2–3), 221–236.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679–698.

- Citti, G., & Sarti, A. (2006). A cortical based model of perceptual completion in the roto-translation space. *Journal of Mathematical Imaging Vision*, 24(3), 307–326.
- Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5), 603–619.
- Cour, T., Benezit, F., & Shi, J. B. (2005). Spectral segmentation with multiscale graph decomposition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (vol. 2, pp. 1124–1131). Piscataway, NJ: IEEE.
- Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 886–893). Piscataway, NJ: IEEE.
- Dollar, P., & Zitnick, C. L. (2013). Structured forests for fast edge detection. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1841–1848). Piscataway, NJ: IEEE.
- Duits, R., & Franken, E. M. (2010a). Left invariant parabolic evolution equations on $SE(2)$ and contour enhancement via invertible orientation scores, part I: Linear left-invariant diffusion equations on $SE(2)$. *Quarterly of Applied Mathematics*, 68, 255–292.
- Duits, R., & Franken, E. M. (2010b). Left invariant parabolic evolution equations on $SE(2)$ and contour enhancement via invertible orientation scores, part II: Nonlinear left-invariant diffusion equations on invertible orientation scores. *Quarterly of Applied Mathematics*, 68, 293–331.
- Felzenszwalb, P. F., & Huttenlocher, D. P. (2004). Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2), 167–181.
- Field, D. J., Golden, J. R., & Hayes, A. J. (2014). Contour integration and the association field. In J. S. Werner & L. M. Chalupa (Eds.), *The new visual neurosciences*. Cambridge, MA: MIT Press.
- Field, D. J., Hayes, A. J., & Hess, R. (1993). Contour integration by the human visual system: Evidence for a local "association field." *Vision Research*, 33, 173–193.
- Forsyth, D., & Ponce, J. (2003). *Computer vision: A modern approach*. Upper Saddle River, NJ: Prentice Hall.
- Ganin, Y., & Lempitsky, V. S. (2014). N4-fields: Neural network nearest neighbor fields for image transforms. In *Proceedings of the Asian Conference on Computer Vision* (pp. 536–551). Singapore: Springer.
- Grigorescu, C., Petkov, N., & Westenberg, M. (2003). Contour detection based on non-classical receptive field inhibition. *IEEE Transactions on Image Processing*, 12(10), 1274–1286.
- Hess, R. F., Hayes, A., & Field, D. J. (2003). Contour integration and cortical processing. *Journal of Physiology-Paris*, 97(2–3), 105–119.
- Hubel, D. H. (1988). *Eye, brain and vision*. New York: Scientific American Library.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive field and functional architecture of monkey striate cortex. *Journal of Physiology*, 195(1), 215–243.
- Hwang, J. J., & Liu, T. L. (2014). Contour detection using cost-sensitive convolutional neural networks. arXiv:1412.6857.
- Isola, P., Zoran, D., Krishnan, D., & Adelson, E. H. (2014). Crisp boundary detection using pointwise mutual information. In *Proceedings of the 13th European Conference on Computer Vision* (pp. 799–814). Zurich, Switzerland.

- Jesse, D., & Mark, G. (2006). The relationship between precision-recall and ROC curves. In *Proceedings of the 23rd International Conference on Machine Learning* (pp. 233–240). New York: ACM.
- Jones, H. E., Grieve, K. L., Wang, W., & Sillito, A. M. (2001). Surround suppression in primate V1. *Journal of Neurophysiology*, 86(4), 2011–2028.
- Kapadia, M. K., Westheimer, G., & Gillert, C. D. (2000). Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *Journal of Neurophysiology*, 84(4), 2048–2062.
- Kivinen, J., Williams, C., & Heess, N. (2014). Visual boundary prediction: A deep neural prediction network and quality dissection. *Journal of Machine Learning Research: Workshop and Conference Proceedings*, 33(1), 512–521.
- La Cara, G. E., & Ursino, M. (2008). A model of contour extraction including multiple scales, flexible inhibition and attention. *Neural Networks*, 21(5), 759–773.
- Landy, M. S., & Graham, N. (2004). Visual perception of texture. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1106–1118). Cambridge, MA: MIT Press.
- Li, C. Y. (1996). Integration field beyond the classical receptive field: Organization and functional properties. *News Physiology Science*, 11(4), 181–186.
- Li, C. Y., & Li, W. (1994). Extensive integration field beyond the classical receptive field of cat's striate cortical neurons: Classification and tuning properties. *Vision Research*, 34(18), 2337–2355.
- Lim, J. J., Zitnick, C. L., & Dollár, P. (2013). Sketch tokens: A learned mid-level representation for contour and object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3158–3165). Piscataway, NJ: IEEE.
- Long, L., & Li, Y. J. (2008). Contour detection based on the property of orientation selective inhibition of non-classical receptive field. In *Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems* (pp. 1002–1006). Piscataway, NJ: IEEE.
- Maire, M., Arbelaez, P., Fowlkes, C., & Malik, J. (2008). Using contours to detect and localize junctions in natural images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–8). Piscataway, NJ: IEEE.
- Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings of the 8th International Conference on Computer Vision* (vol. 2, pp. 416–423). Piscataway, NJ: IEEE.
- Martin, D. R., Fowlkes, C., & Malik, J. (2004). Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 26(5), 530–549.
- Morimichi, N., & Cornelia, F. (2016). *The image torque operator for contour processing*. arXiv:1601.0466.
- Papari, G., & Petkov, N. (2011). An improved model for surround suppression by steerable filters and multilevel inhibition with application to contour detection. *Pattern Recognition*, 44(9), 1999–2007.
- Papari, G., Petkov, P., & Neri, A. (2007). A biologically motivated multiresolution approach to contour detection. *EURASIP Journal on Advances in Signal Processing*, 2007(1), 1–28.
- Petkov, N., & Westenberg, M. A. (2003). Suppression of contour perception by band-limited noise and its relation to non-classical receptive field inhibition. *Biological Cybernetics*, 88(3), 236–246.

- Ren, X. F. (2008). Multi-scale improves boundary detection in natural images. In *Proceedings of the 10th European Conference on Computer Vision* (pp. 533–545). Berlin: Springer.
- Ren, X. F., & Bo, L. (2012). Discriminatively trained sparse code gradients for contour detection. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems*, 25 (pp. 593–601). Red Hook, NY: Curran.
- Sarti, A., & Citti, G. (2015). The constitution of visual perceptual units in the functional architecture of V1. *Journal of Computational Neuroscience*, 38(2), 285–300.
- Sarti, A., Citti, G., & Petitot, J. (2008). The symplectic structure of the primary visual cortex. *Biological Cybernetics*, 98, 33–48.
- Shen, W., Wang, X. G., Wang, Y., & Bai, X. (2015). DeepContour: A deep convolutional feature learned by positive-sharing loss for contour detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3982–3991). Piscataway, NJ: IEEE.
- Tang, Q. L., Sang, N., & Zhang, T. X. (2007). Extraction of salient contours from cluttered scenes. *Pattern Recognition*, 40(11), 3100–3109.
- Ursino, M., & La Cara, G. E. (2004). A model of contextual interactions and contour detection in primary visual cortex. *Neural Networks*, 17(5), 719–735.
- Vilankar, K. P., Golden, J. R., Chandler, D. M., & Field, D. J. (2014). Local edge statistics provide information regarding occlusion and nonocclusion edges in natural scenes. *Journal of Vision*, 14(9), 13–13.
- Walker, G. A., Ohzawa, I., & Freeman, R. D. (2000). Suppression outside the classical cortical receptive field. *Visual Neuroscience*, 17(3), 369–379.
- Wei, H., Lang, B., & Zuo, Q. S. (2013). Contour detection model with multi-scale integration based on non-classical receptive field. *Neurocomputing*, 103, 247–262.
- Xie, S. N., & Tu, Z. W. (2015). Holistically-nested edge detection. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1395–1403). Piscataway, NJ: IEEE.
- 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 Transactions on Image Processing*, 24(8), 2565–2578.
- Yang, K. F., Gao, S. B., Li, C. Y., & Li, Y. J. (2013). Efficient color boundary detection with color-opponent mechanisms. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2810–2817). Piscataway, NJ: IEEE.
- Yang, K. F., Li, C. Y., & Li, Y. J. (2014). Multifeature-based surround inhibition improves contour detection in natural images. *IEEE Transactions on Image Processing*, 23(12), 5020–5032.
- Zeng, C., Li, Y. J., & Li, C. Y. (2011). Center-surround interaction with adaptive inhibition: A computational model for contour detection. *NeuroImage*, 55(1), 49–66.
- Zeng, C., Li, Y. J., Yang, K. F., & Li, C. Y. (2011). Contour detection based on non-classical receptive field model with butterfly-shaped inhibition subregions. *Neurocomputing*, 74(10), 1527–1534.