Signal, Image and Video Processing
https://doi.org/10.1007/s11760-019-01524-2

ORIGINAL PAPER

Multi-scale contour detection model based on fixational eye movement mechanism

Chuan Lin[1] · **Qing Zhang**[1] · **Yijun Cao**[1]

Received: 20 October 2018 / Revised: 7 May 2019 / Accepted: 15 June 2019

Abstract
Physiological evidence has shown that classical receptive field (CRF) responses in the primary visual cortex (V1) can be suppressed by its surrounding region, called the non-classical receptive field (nCRF). Currently, the contour detection model based on the physiological characteristics of the V1 region is mainly used to suppress texture and highlight contour information through the inhibition of nCRF features. However, the effect of eye movement on inhibition is not considered in the inhibition calculation of such models. Inspired by the fixational eye movement (FEyeM) mechanism, we propose a multi-scale contour detection model based on fixational eye movement (MsFem) and the surrounding suppression mechanism. A bank of filters was proposed to simulate the influence of FEyeMs on nCRF, and multi-scale cues were utilized to improve the fine and coarse contour extraction and texture inhibition. The experiments showed that MsFem outperformed some biologically motivated ones in retaining the small-scale target contour information and suppressing the large-scale background textures.

Keywords Contour detection · Non-classical receptive field · Multi-scale integration · Fixational eye movement

1 Introduction

Contour detection is a fundamental task in the fields of machine vision and target recognition [1, 2]. Early edge detectors were mainly based on a measure of discontinuity in image features, such as the Sobel and Prewitt operators [3]. Compared to edge detection, contour detection focuses on suppressing the background textures and highlights the edge of the target object.

Several contours or edge detection methods have been proposed. Typical examples include local gradient [3], statistical measures [4], graph-based methods [5], region-based methods [6], and biological measures [8–14]. However, there is still room for improvement in existing algorithms and the development of novel and efficient models.

The human visual system (HVS) is an efficient encoded network that can detect salient contours. Neurophysiological and physiological studies [7] have found that the responses of the classical receptive field (CRF) can be suppressed by its surrounding non-CRF (nCRF) in the primary visual cortex (V1). Inspired by neural properties, in recent decades, many biologically motivated contour detection models have been proposed. These methods have shown good performance for grayscale natural images. Grigorescu et al. [8] proposed two inhibition models (i.e., the anisotropic inhibition model and isotropic model) that utilize a two-dimensional (2D) Gabor function to simulate the spatial properties of the CRF and a 2D Gaussian difference (DoG) function to simulate the spatial properties of nCRF. Sang et al. [9, 10] proposed a butterfly-shaped function to modulate the suppression region and a dynamic feedback process to detect the target contour. Huang et al. [11] considered the difference between the center and the stimulation of its neighborhood, which can effectively balance the enhancement and inhibition of neurons. The proposed model incorporates two mechanisms of contextual modulation, namely surround suppression and collinear facilitation. Yang et al. [12] and Han et al. [13] further proposed a model based on multiple features, which revealed the potential role of the interaction between model V1 neurons and the direction. It is revealed that surrounding inhibition combining multiple local cues can significantly improve contour detection in complex scenes. Xiao et al. [14] proposed a model using the combined contextual modulation to change the computational complexity and optimize the detection

✉ Chuan Lin
chuanlin@gxust.edu.cn

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

Published online: 20 July 2019

results. For color images, Yang et al. [15] proposed a CO model to detect brightness edges and color edges based on a color opponent mechanism. Subsequently an SCO model was proposed by introducing a texture sparseness constraint [16]. Akbarinia et al. [17] proposed a model that includes four kinds of receptive field surround and uses the V2 to integrate single-channel information.

The above research, based on the vision mechanism contour detection method, was conducted for paralyzed animals; thus, fixational eye movement (FEyeM) was not considered. FEyeMs refer to the tiny unconscious movements of the eye in a fixation state; they play an important role in HVS. It is possible to correct the positional shift caused by quiver and drift and adjust the retina image to the optimal state [18–20]. In addition, the existing literature does not specify the relationship between FEyeM and computational models. Inspired by the FEyeM mechanism, we propose a model to compute nCRF responses, which can efficiently suppress textures.

The multi-scale analysis tool is very important in image processing [24–28] and image super-resolution reconstruction [21–23]. Some models consider the multi-scale information of the image to introduce further developments in the field. Zeng et al. [24, 25] calculated neuron inhibition with two scales, where the small scale was suppressed on two sides and the large scale was enhanced on two ends of the butterfly shape inhibition function. The butterfly-shaped model only activates one inhibitory subregion and is based on orientation saliency characteristics. Wei et al. [26] not only proposed a multi-scale integration model that computed the scale of each neuron according to its dynamic properties, but also introduced preprocessing of the external information by the retinal ganglion cell at an early stage of visual processing. Yang et al. [27] proposed a multi-feature-based model that combined multiple suppression weights, such as orientation, luminance, and luminance contrast; then, the combined weights are used to modulate the final surround inhibition of the neurons. Cai et al. [28] proposed a new multi-scale contour detection model using deep neural network technology. Although the multi-scale contour detection has achieved good results, there are still some problems to be solved. When the Gaussian function is used for filtering, the contour will become smooth as the receptive field increases in scale; however, the contour will generate a certain amount of contour information deviation, compared with the small scale.

The main contributions of this include the following. (1) We used twelve sub-templates within equal intervals to modulate the regional displacement of neurons during the FEyeM mechanism and then obtained the final suppression term by choosing the minimum value from the twelve sub-template responses. (2) Inspired by the multi-level synchronization information processing mechanism in the visual cortex, the multi-scale space theory is incorporated into the proposed eye movement inhibition template. Combined with multi-scale analysis, a multi-scale fusion method is proposed to solve the problem of information deviation effectively.

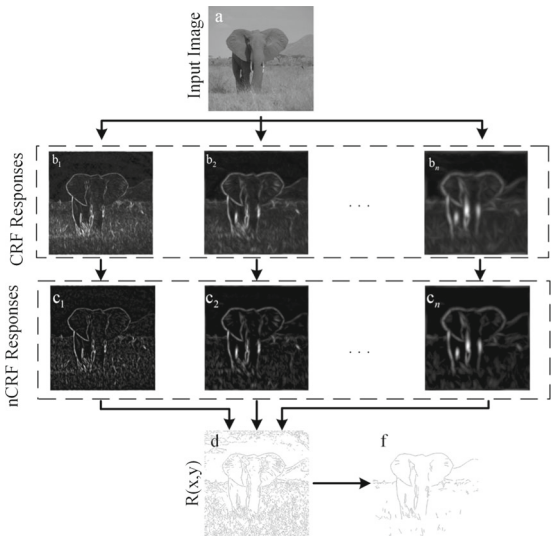

Fig. 1 Flowchart of the proposed contour detection model in this paper. (a) Input image. (b_1–b_n) Maximum Gabor energy maps on different scales. (c_1–c_n) Energy maps after inhibition. (d) Contour map after multi-scale integration. (f) Contour map after binarization

2 Computational models

2.1 Overview

A good contour detection model should produce greater inhibition in the regions containing texture information. For the same image, there is a trade-off between removing the maximum amount of the background texture and retaining contour information. To better solve this problem, we propose a method that uses multi-scale integration to obtain the final contour map. Figure 1 shows a flowchart of the proposed model. From this flowchart, CRF responses will be computed by a grayscale image convolution with a bank of Gabor filters. The corresponding nCRF responses will suppress the CRF responses, as shown in Fig. 1(c_1–c_n). The fused image $R(x, y)$ will process further non-maximal suppression and hysteresis threshold processing; subsequently, a binarized contour image is obtained as the output. In addition, we use the derivative of Gaussian (GD) function to simulate the characteristics of the CRF in V1 and obtain good results.

2.2 Responses of CRF

2.2.1 Gabor filter and GD filter

A 2D Gabor filter describes the spatial properties of the simple cells in V1 [8], which is a Gaussian multiplied by a cosine-modulated wave.

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

$$\tilde{x} = x\cos\theta + y\sin\theta, \quad \tilde{y} = -x\sin\theta + y\cos\theta. \quad (2)$$

The λ parameter is the wavelength of the Gabor filter function, $1/\lambda$ is the cosine function of the spatial frequency, and σ/λ represents the spatial frequency bandwidth. Referring to the work of Grigorescu et al. [8], we set $\sigma/\lambda = 0.56$. The coefficient σ is the standard deviation of the Gaussian factor, which determines the size of the Gabor filter. Θ is the direction of the CRF parameters. φ represents the phase parameter. $\gamma = 0.5$ is a spatial aspect ratio representing the ratio between the long and short axes of the elliptical receptive field.

In a recent study [27], the GD function was used to simulate the characteristics of the CRF in V1. The GD function can be defined as follows:

$$GD(x, y; \sigma, \theta) = \frac{\partial G(\tilde{x}, \tilde{y}; \sigma, \theta)}{\partial\tilde{x}}, \quad (3)$$

$$G(\tilde{x}, \tilde{y}; \sigma, \theta) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{\tilde{x}^2 + \gamma\tilde{y}^2}{2\sigma^2}\right), \quad (4)$$

where θ is the orientation of the CRF parameters. $\tilde{x}$ and $\tilde{y}$ are calculated using (Eq. 2). We set spatial aspect ratio $\gamma = 0.5$ [27]. The standard deviation σ determines the size of the receptive field.

2.2.2 Energy map construction

Simple cell responses have a strong selective orientation. Therefore, a Gabor filter function with multiple orientations must be preset to find the optimal orientation of the receptive field. From [8], the response of simple cells $e(x, y; \sigma, \theta, \varphi)$ at different orientations are obtained by convoluting the original input image $h(x, y)$ with the Gabor filter function $g(x, y; \sigma, \theta, \varphi)$. Therefore, the response of simple cells can be defined as follows:

$$e(x, y; \sigma, \theta, \varphi) = h(x, y) * g(x, y; \sigma, \theta, \varphi), \quad (5)$$

where * is the convolution operator.

The complex cell receptive response $E(x, y; \sigma, \theta)$ is composed of a pair of simple cell responses with a phase difference of $\pi/2$:

$$E(x, y; \sigma, \theta) = \sqrt{e^2(x, y; \sigma, \theta, 0) + e^2\left(x, y; \sigma, \theta, -\frac{\pi}{2}\right)}. \quad (6)$$

From [27], another CRF response of a V1 cell $E = (x, y; \sigma, \theta)$ calculated by a GD filter is defined as:

$$E(x, y; \sigma, \theta) = h(x, y) * GD(x, y; \sigma, \theta). \quad (7)$$

Then, we obtain Gabor and GD energy maps $E\left(x, y; \sigma, \theta_j\right)$ in different directions. θ_j is defined as follows:

$$\theta_j = \frac{(j-1)\pi}{N_\theta}, j = 1, 2, \ldots, N_\theta, \quad (8)$$

where we set $N_\theta = 12$.

The Gabor filter and GD filter have clear orientation selectivity. Each neuron presents an optimal response along the optimal orientation. The maximum energy map $\hat{E}(x, y; \sigma)$ and corresponding orientation $\emptyset^I(x, y)$ are shown below:

$$\hat{E}(x, y; \sigma) = \max\left\{E\left(x, y; \sigma, \theta_j\right) | j = 1, 2, \ldots, N_\theta\right\}, \quad (9)$$

$$k = \arg\max\left\{E_{\sigma,\theta_j}(x, y) | j = 1, 2, \ldots, N_\theta\right\}, \quad (10)$$

$$\emptyset^I(x, y) = \theta_k. \quad (11)$$

2.3 Responses of nCRF

Inspired by the FEyeM mechanism, we present a multi-scale nCRF suppression template based on the autonomic healing characteristics of the receptive fields and the multi-directional movement competition. A new template is proposed based on the DoG inhibition template mentioned in [8], which is simulated using a uniformly weighted function over the surrounded nCRF circle. Simultaneously, the multi-scale information channels in the visual cortex are simulated by incorporating multiple scaling parameters.

2.3.1 FEyeM mechanisms

Physiological studies have found that unconscious eye movements prevent visual perception from completely disappearing as the neuron adapts to its invariant environment. The visual system during fixation of the main eye movement is divided into high-frequency tremors, drift, and flashing movements [29].

The physiological properties of fixation movement studies show that the flashing movement can be used to repair displacement deviations due to tremor or drift. This means

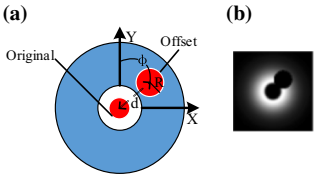

Fig. 2 **a** Physiological phenomenon of FEyeMs. **b** computational kernel $W_{\mathrm{d}}(x, y; \sigma, \phi)$ of FEyeM model

that the receptive field of the FEyeM mechanism has its own characteristics of recovering. To make the contour detection model in line with the physiological characteristics of the receptive field, we propose a new function, as shown in Fig. 2, to simulate the processes of FEyeMs.

High-frequency tremor and drift make retinal images deviate from the center position and then repair them by the flashing movement, as shown in Fig. 2a. During computation, we divide the nCRF region equally into twelve different directional movement models to simulate the random movement of FEyeMs. The offset position in the FEyeM template can be modeled by a function of the form $U_r(x, y; d, \phi)$. Then, we obtain $U_r\left(x, y; d, \phi_j\right)$ in different positions. ϕ_j is defined by Eq. (13)

$$U_r(x, y; d, \phi) = \begin{cases} 0, (x - d\sin\phi)^2 + (y - d\cos\phi)^2) \le R^2 \\ 1, (x - d\sin\phi)^2 + (y - d\cos\phi)^2) > R^2 \end{cases}, \tag{12}$$

$$\phi_j = \frac{\pi(j-1)}{N_\phi}, j = 1, 2, \ldots, N_\phi, \tag{13}$$

where N_ϕ is the number of movement templates; we set $N_\phi = 12$ in this experiment, and R is the radius of the offset area. As shown in Fig. 2a, d in $U_r(x, y; d, \phi)$ represents the distance between the center of the original location and the offset position center. We set $R = 3\sigma$ and $d = 5\sigma$. ϕ is the offset angle, which is the angle between the center of the offset position and the positive semi-axis of the y-axis.

Each sub-template function $S(x, y; \sigma, \phi)$ is multiplied by a 2D DoG function and $U_r(x, y; d, \phi)$ function, as shown in Eq. (15). A 2D DoG function is shown in Eq. (16). The inhibitory property of simulated cell nCRF is modeled by normalization:

$$W_{\mathrm{d}}(x, y; \sigma, \phi) = \frac{N(S(x, y; \sigma, \phi))}{\left\| N(S(x, y; \sigma, \phi))_1 \right\|}, \tag{14}$$

$$S(x, y; \sigma, \phi) = U_r(x, y; d, \phi) * \mathrm{DoG}(x, y; \sigma), \tag{15}$$

$$\mathrm{DoG}(x, y; \sigma) = \frac{1}{2\pi(4\sigma)^2}\exp\left(-\frac{x^2+y^2}{2(4\sigma)^2}\right) - \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \tag{16}$$

where $\|*\|_1$ is the L1 normal form. The function $N(x) = \max(0, x)$ is a half-wave rectification operation, indicating that an activated neuron only received a positive value.

The neural responses $F_i(x, y)$ of V1 region are composed of the CRF stimulation $\hat{E}(x, y; \sigma_i)$ and nCRF inhibition $\mathrm{Inhib}(x, y)$, as shown in Eq. (17). The suppression response $\widetilde{Inh}(x, y; \phi)$ of the nCRF at each scale is obtained by convolution of the energy map and the distance weight function $W_{\mathrm{d}}(x, y; \sigma, \phi)$ at the corresponding scale, as shown in Eq. (19). Then, the minimum value from the suppression of the twelve orientations is chosen as the optimal suppression response at the scale, as shown in Eq. (18).

$$F_i(x, y) = N\left(\hat{E}(x, y; \sigma_i) - \alpha Inhib(x, y)\right), \sigma_i \in [1, 2, 4, 8 | i = 1, 2, 3, 4], \tag{17}$$

$$Inhib(x, y) = \min(\widetilde{Inh}(x, y; \phi_j) | j = 1, 2, \ldots, 12), \tag{18}$$

$$\widetilde{Inh}(x, y; \phi) = \hat{E}(x, y; \sigma) * W_{\mathrm{d}}(x, y; \sigma, \phi), \tag{19}$$

where α controls the weight of the inhibition of the surrounding neurons.

2.3.2 Multi-scale integration

The multi-scale of the physical scene is reflected by the scale changes in the cell's receptive field. The role of multi-scale analysis in neurophysiological studies is becoming increasingly important.

Based on the scale space theory, each scale value corresponds to the size of a set of neuron receptive fields. Selecting the appropriate scale can achieve preservation of the contour and elimination of the texture. Many current contour detection methods only consider the improvement in the nCRF inhibition on a single scale. With multi-scale information fusion, the texture background can be removed significantly on the premise of maintaining the integrity of the contour, which is in line with the spatial characteristics of the visual receptive field.

Contour information will shift as the receptive field (RF) scale increases, which is illustrated by example with zoomed-in local regions shown in Fig. 3. The red line represents the contour of the area, and the blue region marks the offset of the contour. Figure 3 shows that the contour offset increases as the RF scale is increased.

In order to solve the problem of contour information deviation at a large scale, we propose a multi-scale suppression method. Figure 4 shows the process of the proposed multi-scale inhibition method. First, the responses of different scales are reduced using the non-maximum suppression (NMS) followed hysteresis thresholding. Based on the binary map of lowest scale, we gradually compute the responses at

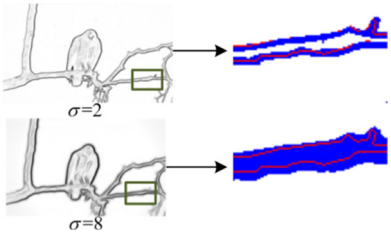

Fig. 3 Contour information will shift with the increase in RF scale

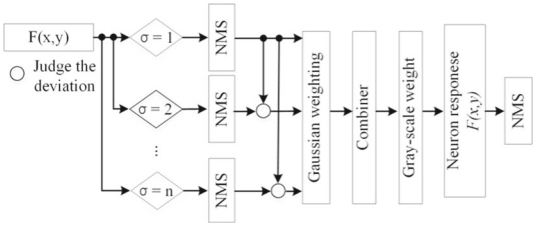

Fig. 4 Process of the proposed multi-scale inhibition model

each high scale, which is considered in this computational process in addition to the localization in different scales. To obtain the responses and localization information in different scales, we set a sliding window, has have different radii increasing along with the increased scale, at each contour pixel. The sliding window is of the form given below:

$$M(x, y; \sigma, \beta) = \begin{cases} 0, & |\beta| \leq \frac{\pi}{4} \\ 1, & \text{otherwise} \end{cases}, \tag{20}$$

where β represents the angle between the preferred orientation of CRF. In this formula, σ is the scale parameter used in the Gabor function (Eq. 1), which controls the radius of the sliding windows $r = \sigma - \sigma_1$, where σ_1 is the σ corresponding to the lowest scale.

In order to determine the deviation information of pixels at different scales apart from the lowest one, we define the deviation information response $L(x, y, \sigma, i)$ as the convolution of the binary contour map $\bar{F}_i(x, y)$ with the sliding window $M(x, y; \sigma, \beta)$:

$$L(x, y, \sigma, i) = \bar{F}_i(x, y) * M(x, y; \sigma, \beta), \tag{21}$$

where $\bar{F}_i(x, y)$ is the responses of $F_i(x, y)$ as (Eq. 17) after NMS and hysteresis thresholding.

We then weighted the resulting deviation information response. The larger the weight, the more important the contour deviation information at this scale. As the scale increases, the deviation of the contour pixel point increases. To assign weights to the deviation information, a decreasing function is selected to fit the deviation information across different scales; therefore, a Gaussian is used. The contour map $T(x, y)$ of the multi-scale fusion can be obtained by summing the weighted deviation information:

$$T(x, y) = \left(\bar{F}_1(x, y) + \sum_{i=1}^{N_j} L(x, y, \sigma_i, i) \right) * \frac{1}{\sqrt{2\pi}\varepsilon} \exp\left(-\frac{(\sigma_i - \mu)^2}{2\varepsilon^2} \right), \sigma_i \in [1, 2, 4, 8], \tag{22}$$

where N_j is the number of σ_i; $\varepsilon = 3$, $\mu = 1$. We resume the steps of our multi-scale suppression method in Algorithm 1.

We obtain the number of pixels corresponding to each gray scale in the contour map $T(x, y)$ except 0 as the weight matrix. m_j indicates that there are m pixels in the j gray scale, and M denotes the total number of pixels in the gray scale. Then, the weight of each gray scale is $(1 + m_j/M)$. The gray scale of each pixel of neuron response $F_i(x, y)$ is calculated and multiplied by the corresponding grayscale weight. We perform the NMS and hysteresis threshold processing on the newly obtained neuron response $R(x, y)$, and take the optimal contour map in the scale as the contour map in this study.

Algorithm 1 *Multi-scale suppression algorithm.*

Input: scale $\sigma_i \in \{1,2,4,8\}$; binary contour map $\bar{F}_i(x, y)$, i ∈ {1,2,3,4}; β = π/4
Begin
for i ← 2 to 4 **do**
  $M(x, y; \sigma_i, \beta)$ ← compute eq. (20);
  for each pixel $(x, y) \in \bar{F}_i(x, y)$ **do**
    $L(x, y, \sigma_i, i)$ ← compute eq. (21);
    if (x, y) in $\bar{F}_i(x, y)$ is not deviation pixel
      pixel (x, y) in $L(x, y, \sigma_i, i)$ ← l;
    else
      pixel (x, y) in $L(x, y, \sigma_i, i)$ ← 0;
  end for
end for
for (x, y) in $L(x, y, \sigma_i, i)$
  $hit(x, y)$ ← zeros(size($L(x, y, \sigma_i, i)$));
  for i ← 2 to 4 **do**
    $temp(x, y)$ ← $L(x, y, \sigma_i, i)$ * Gaussian(ε = 3, μ = 1, σ_i) //Gaussian function as eq. (22)
    $hit(x, y)$ ← $temp(x, y)$ + $hit(x, y)$;
  end for
end for
$T(x, y)$ ← **sum**($\bar{F}_1(x, y)$, $hit(x, y)$)
End

3 Experimental results

In order to verify the effectiveness and feasibility of the MsFem model, we use the RuG dataset [8] and Berkley Segmentation Dataset (BSDS) [27].

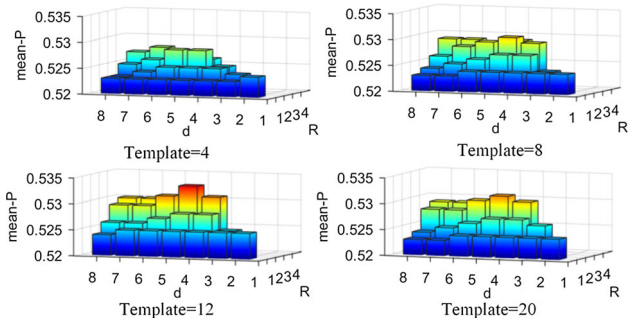

Fig. 5 Mean performance over all the images of the RuG40 dataset with various parameter combinations

Table 1 Comparison of the quantitative performance of various models

Model	ODS	OIS	Best P
MCI [27]	0.47	0.49	0.52
LI [27]	0.45	0.49	0.50
OI [27]	0.44	0.48	0.49
CI [27]	0.43	0.49	0.50
SCO [16]	0.46	0.51	0.52
SED [17]	0.45	0.49	0.49
MsFem-Gabor w/o FEyeMs	0.46	0.50	0.52
MsFem-Gabor w/o multi-scale	0.44	0.47	0.48
MsFem-Gabor	0.47	0.51	0.53
MsFem-GD	0.47	0.52	0.54

3.1 Model performance analysis

In the FEyeM model, several sets of parameters R and d were selected to simulate their physiological characteristics, where the parameters R and d are set as the parameters in Fig. 2a. Figure 5 shows the average of P over all the images of the entire RuG dataset with different parameters. Each set of parameters combines the parameter R and parameter d, where $R < d$. Figure 5 depicts the result of the selection using Gabor filter. We set four sets of movement template parameters, where the template parameters are obtained by $N_\phi = \{4, 8, 12, 20\}$ in Eq. (13). Figure 5 shows that when the number of templates N_ϕ is 12, a higher average P is obtained; when $R = 3\sigma$, $d = 5\sigma$, the average P is the largest. Therefore, we select $N_\phi = 12$, $R = 3\sigma$, and $d = 5\sigma$ as optimal parameters.

Through the experimental analysis, contour detection in this work is more complete and suppresses more texture information. The advantage of the multi-scale method over the single-scale method is that the detected contour information is more detailed for small-scale parameters. Although contour detection is rough under large-scale parameters, the backgrounds are relatively small. Combining the two can lead to the preservation of the fine outline at small scales and reduce the texture background. Thus, the use of FEyeMs can describe the physiological characteristics of the eye; when the eye is looking at a stationary object, FEyeMs can highlight the location of the edge.

3.2 Experiment on RuG dataset

In this paper, the evaluation criteria proposed by Grigorescu et al. were used to evaluate the detection performance. Precision P is defined as the fraction of true contour pixels from all detected pixels, where $P \in [0, 1]$. T_{GT}, and T_{GF} are the edge pixels and background texture pixels of the truth contour image, respectively. T_{DT} and T_{DF} are the image contour pixels and the background pixels set in the final detection result, respectively. Thus, we define the correctly detected contour pixels as $T = T_{GT} \cap T_{DT}$, falsely detected contour pixels as $T_{FP} = T_{DT} \cap T_{GF}$, and missed contour pixels as $T_{FN} = T_{DF} \cap T_{GT}$. The performance evaluation criteria P, false detection rate E_{FP}, and missed detection rate E_{FN} are as follows:

$$P = \frac{\text{card}(T)}{\text{card}(T) + \text{card}(T_{FP}) + \text{card}(T_{FN})}, \tag{23}$$

where card (Y) represents the number of set elements in set Y. As the contour cannot be detected as an exact integer image coordinate every time, we used a 5×5 square region of tolerance when matching a contour pixel in the model output to the ground-truth contours.

In order to test the proposed model, we compared it with the state-of-the-art methods currently based on gray contour detection [27]. Table 1 shows a comparison of the quantitative performance of three models, namely optimal dataset scale (ODS), optimal image scale (OIS), and best precision (Best P). In [27], the LI model is based on luminance suppression cues, the OI model is based on orientation suppression cues, the CI model is based on luminance contrast suppression cues, and the MCI model is their fusion, which is a multi-cue suppression model. However, the model proposed in this paper considers only distance weight suppression. Moreover, our experimental results are better than those obtained using the compared models.

3.3 Experiment on BSDS

We tested MsFem on the BSDS dataset [27]. In this experiment, we used the precision–recall (PR) curve mentioned in the BSDS and its F-score to measure the performance. That is, contour detection is considered as a classification problem herein to evaluate the test results. The PR framework is used to evaluate the accuracy of the contour detection model by

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

Model	BSDS300			BSDS500		
	ODS	OIS	AP	ODS	OIS	AP
LI [27]	0.59	0.63	0.60	0.61	0.63	0.60
OI [27]	0.59	0.63	0.60	0.61	0.63	0.60
CO [15]	0.60	0.63	0.60	0.61	0.64	0.61
SCO [16]	0.62	0.64	0.64	0.63	0.67	0.66
SED [17]	0.65	0.67	0.68	0.67	0.70	0.70
MsFem-GD	0.61	0.63	0.60	0.62	0.64	0.61
MsFem-Gabor	0.61	0.63	0.56	0.62	0.64	0.57

using the manually drawn contour marked on the BSDS as a standard contour. The F-score is defined as:

$$F = 2PR/(R + P), \tag{24}$$

where $P = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP})$, $R = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$ P is the precision, that is, the measure of exactness of the detected edges. R is the recall, which is the coverage of the contour in the real contour. TP, FP, and FN represent the number of true positives, false positives, and false negatives, respectively.

The BSDS contains two sets of images BSDS300 (100 test images and 200 training images) and BSDS500 (200 test images). Quantitative comparison was also conducted using BSDS300 and BSDS500. We used the same parameter settings as those on the RuG dataset, with the exception of p, owing to which hysteresis thresholding was not used in this experiment. We selected the OI inhibition model (OI model), LI inhibition model (LI model), CO inhibition model [15] (CO model), MsFem-GD, and MsFem-Gabor for comparison. Note that we used tolerant parameters maxDist = 0.0075 in the BSDS benchmark and grayscale version in BSDS images. Table 2 compares the quantitative performance of more models, which includes the ODS, OIS, and average precision (AP). Figure 6 shows the different contour detection model PR curves and F-score for BSDS300. Figure 7 shows the results of contour detection on four natural images from BSDS300.

The proposed model is lower than the SCO and SED on BSDS, while it obtains better performance on RuG dataset due to the fact that different models have distinctive advantages and are suitable for different datasets. Note that SCO and SED use extra feature to extract contours comparing to proposed model. Specifically, we only use physiological characteristics of V1 region, whereas SCO uses spatial sparseness constraint and SED uses physiological characteristics of V2 neurons.

Through the above analysis of image data in the BSDS300 dataset, we can see that the MsFem model has a higher F-score compared to the other detection models mentioned in this paper. The MsFem contour detection method has high accuracy, yields good performance, and is in line with the physiological characteristics of the perception mechanism.

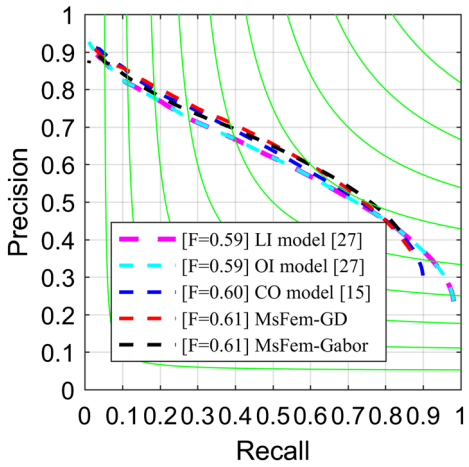

Fig. 6 PR curve (and F-scores) of the different contour detection models for BSDS300

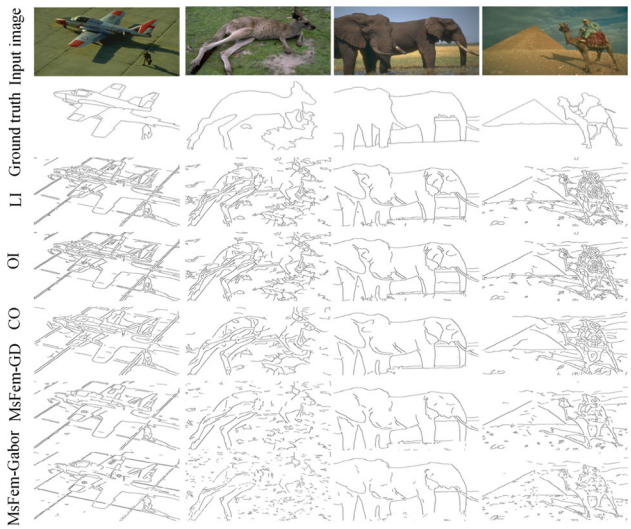

Fig. 7 Test results of randomly selected four images from BSDS300. The first two last lines represent input images, ground-truth contour, LI model, OI model, CO model, MsFem-GD, and MsFem–Gabor

4 Discussion and conclusions

The contour detection model presented in this paper has two distinct aspects compared to the existing points of innovation of the contour detection model.

Based on the contour detection of the traditional receptive field, the introduction of FEyeMs brings contour detection in line with the physiological characteristics. In the traditional DoG template, we used twelve sub-templates within equal intervals. Then, the corresponding angular orientation area

in the sub-template is exchanged with the central area of the template to obtain the final suppression term by integrating the weights.

As the amplitude of eyeball varies with the actual vibration, several groups of movement-receptive field-scale models with different scales are introduced in this paper. The size of the receptive field is related to the radius of the offset position in the movement suppressor sub-template. The multi-scale contour detection model retains the small-scale target contour information and suppresses the large-scale background texture so that the accuracy of contour detection can be further improved.

In this study, we used the Gabor filter and GD filter to test the RUG dataset and BSDS300, respectively. The experiments showed that the results of the RUG dataset were not as good as those of the GD filter when the Gabor filter was used. However, in BSDS300, Gabor filter's F-score was the same as that of the GD filters. The Gabor energy is a combined result that includes even- and odd-symmetric filtering outputs. The GD filter employs a process of performing partial derivative and Gaussian filtering on an image. The odd-symmetric filter can detect the difference in brightness between the two sides, and the even-symmetric filter can detect the difference between the center and the surrounding brightness. Gabor filter trade-offs may omit some contours, while the GD filter retains more contour information; however, it will introduce more noise in the detection results than the Gabor filter. Therefore, different models have different performances in different image datasets.

Subsequent improvements can be considered from the following aspects. (1) To determine whether the position, shape, and range of the offset region can be changed will require a large number of experiments; only a preliminary scheme is proposed in this paper. (2) To determine whether there is a better method to combine the contour information of multiple scales in the process of multi-scale fusion requires further research.

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

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interests regarding the publication of this paper.

References

1. Davies, E.R.: Machine Vision: Theory, Algorithms, Practicalities. Elsevier, Amsterdam (2004)
2. Milan, S., Roger, B., Vaclav, H.: Image processing, analysis, and machine vision. J. Electron. Imaging **9**(82), 685–686 (2014)
3. Baglodi, V.: Edge detection comparison study and discussion of a new methodology. In: Southeastcon, Southeastcon 09 IEEE (2009)
4. Dong, H.L., Jang, S.J.: Comparison of two-sample tests for edge detection in noisy images. J. R. Stat. Soc. **51**(1), 21–30 (2002)
5. Nouri, F., Kazemi, K., Danyali, H.: Salient object detection using local, global and high contrast graphs. Signal Image Video Process. **12**(4), 659–667 (2018)
6. Yadollahi, M., Procházka, A., Kašparová, M., Vyšata, O.: The use of combined illumination in segmentation of orthodontic bodies. Signal Image Video Process. **9**(1), 1–8 (2014)
7. Tang, Q., Sang, N., Zhang, T.: Extraction of salient contours from cluttered scenes. Pattern Recognit. **40**(11), 3100–3109 (2007)
8. Grigorescu, C., Petkov, N., Westenberg, M.A.: Contour detection based on nonclassical receptive field inhibition. IEEE Trans. Image Process. **12**(7), 729–739 (2003)
9. Nong, S., Li, H., Peng, W., Zhang, T.: Knowledge-based adaptive thresholding segmentation of digital subtraction angiography images. Image Vis. Comput. **25**(8), 1263–1270 (2007)
10. Nong, S., Qiling, T., Tianxu, Z.: Contour detection based on inhibition of primary visual cortex. J. Infrared Millim. Waves **26**(1), 47–51 (2007)
11. Huang, W., Jiao, L., Jia, J., Yu, H.: A neural contextual model for detecting perceptually salient contours. Pattern Recognit. Lett. **30**(11), 985–993 (2009)
12. Yang, K., Li, Y.: A coutour detection model based on surround inhibition with multiple cues. In: Chinese Conference on Pattern Recognition (2012)
13. Jing, H., Jiang, Y., Yi, Z., Bai, L.F.: Salient contour extraction from complex natural scene in night vision image. Infrared Phys. Technol. **63**(11), 165–177 (2014)
14. Xiao, J., Cai, C.: Contour detection based on horizontal interactions in primary visual cortex. Electron. Lett. **50**(5), 359–361 (2014)
15. Yang, K., Gao, S., Li, C., Li, Y.: Efficient color boundary detection with color-opponent mechanisms. In: Computer Vision & Pattern Recognition (2013)
16. Yang, K., Gao, S., Guo, C., Li, C., Li, Y.: Boundary detection using double-opponency and spatial sparseness constraint. IEEE Trans. Image Process. **24**(8), 2565–2578 (2015)
17. Akbarinia, A., Parraga, C.A.: Biologically-inspired edge detection through surround modulation. In: Proceedings of the British Machine Vision Conference, pp. 1–13 (2016)
18. Costela, F.M., McCamy, M.B., Macknik, S.L., Otero-Millan, J., Martinez-Conde, S.: Microsaccades restore the visibility of minute foveal targets. PeerJ **1**, e119 (2013)
19. Sui, X., Hang, G., Sun, Y., Qian, C., Gu, G.: Infrared super-resolution imaging method based on retina micro-motion. Infrared Phys. Technol. **60**(5), 340–345 (2013)
20. Martinez-Conde, S., Otero-Millan, J., Macknik, S.L.: The impact of microsaccades on vision: towards a unified theory of saccadic function. Nat. Rev. Neurosci. **14**(2), 83–96 (2013)
21. Wei, D.: Image super-resolution reconstruction using the high-order derivative interpolation associated with fractional filter functions. IET Signal Proc. **10**(9), 1052–1061 (2017)
22. Wei, D., Li, Y.M.: Generalized Sampling Expansions with Multiple Sampling Rates for Lowpass and Bandpass Signals in the Fractional Fourier Transform Domain. IEEE Trans. Signal Process. **64**(18), 4861–4874 (2016)

23. Wei, D., Li, Y.: Reconstruction of multidimensional bandlimited signals from multichannel samples in linear canonical transform domain. Signal Process. IET **8**(6), 647–657 (2014)
24. Zeng, C., Li, Y., Yang, K., Li, C.: Contour detection based on a non-classical receptive field model with butterfly-shaped inhibition subregions. Neurocomputing **74**(10), 1527–1534 (2011)
25. Zeng, C., Li, Y., Li, C.: Center–surround interaction with adaptive inhibition: a computational model for contour detection. NeuroImage **55**(1), 49–66 (2011)
26. Wei, H., Lang, B., Zuo, Q.: Contour detection model with multi-scale integration based on non-classical receptive field. Neurocomputing **103**, 247–262 (2013)
27. Yang, K.-F., Li, C.-Y., Li, Y.-J.: Multifeature-based surround inhibition improves contour detection in natural images. IEEE Trans. Image Process. **23**(12), 5020–5032 (2014)
28. Cai, Z., Fan, Q., Feris, R.S., Vasconcelos, N.: A unified multi-scale deep convolutional neural network for fast object detection. In: European Conference on Computer Vision, pp. 354–370. Springer (2016)
29. Martinez-Conde, S., Macknik, S.L., Hubel, D.H.: The role of fixational eye movements in visual perception. Nat. Rev. Neurosci. **5**(3), 229–240 (2004)