



A contour perception model that simulates the complex connection pattern of the visual cortex

Zhefei Cai¹ · Yingle Fan¹

Received: 22 September 2021 / Revised: 20 April 2022 / Accepted: 27 October 2022 /
Published online: 21 November 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Contour detection is the basic content of image processing and plays an important role in image analysis and target recognition. This paper proposed a contour perception model that simulates the complex connection pattern of the visual cortex. The connection included the feedforward input from the lateral geniculate body (LGN), the horizontal input from the neurons in the same layer, and the feedback input from the advanced visual cortex. Using the sparse coding characteristics of the LGN, the windmill-like structure receptive field of the primary visual cortex, and the hue perception characteristics of the advanced visual cortex to improve the accuracy of the contour extracted by the proposed model. Choosing the BSDS500 natural scene dataset as the experimental object, the F-score is selected as the evaluation index. The average optimal F-score of the proposed method is 0.72, which is better than other mainstream biological vision-based methods. Concurrently, the NYUD dataset is used for further verification. To comprehensively verify the effectiveness of the model proposed in this paper, Performance-value rather than F-score is selected as the evaluation index. The average optimal Performance-value of the proposed method is 0.42, which shows better results, too. The complex connection pattern allows neural encoding and decoding to make full use of the characteristics of information exchange between the visual cortexes, which is more in line with the biological vision system.

Keywords Feedback connection · Horizontal connection · Hue perception · Surround suppression

1 Introduction

Contour refers to the lines that constitute the outer edge of the image or object. It not only helps to intuitively understand the information contained in the image but also affects the subsequent

✉ Yingle Fan
fan@hdu.edu.cn

¹ Laboratory of Pattern Recognition and Image Processing, Hangzhou Dianzi University, Hangzhou 310018, China

analysis and processing of the image. Contour detection is the basic content of image processing. It aims to extract the target contour as completely as possible while removing the background and internal texture noise. It was used in computer vision such as semantic edge detection [21], semantic segmentation [30], and image repair [39].

Contour detection methods in the early time mainly used the low-level features of the image, relying on the brightness, color, and contrast to judge the difference between adjacent pixels, such as the Canny algorithm [7], Pb algorithm [22], and so on. Pixel-level contour detection methods are insufficient in accuracy. Therefore, some researchers have proposed the methods of contour detection which used the middle-level features in the image, that is, local features, such as Sketch Token [16] method and Structured Forests for Fast Edge Detection(SE) [10] method. With the popularity of deep learning methods, contour detection models pay more attention to the high-level abstract features of the image. The DeepEdge model [5] adopted the local high-level features, using a multi-scale deep network composed of a local feature extraction network and a fully connected network to extract contours. The HED model [32] used the global high-level features to achieve the contour detection task of end-to-end training. It was the first time that end-to-end training has been successfully used for contour detection tasks. Combining saliency detection and contour detection, the MLM model [31] improved the accuracy of contour detection and saliency target boundary detection at the same time. Also, using a generative adversarial network(GAN) to improve the contour quality was endless [19, 35, 38].

Thanks to the continuous development of biological vision research, various vision mechanisms have been applied to image processing. Based on the biological vision, most of the contour extraction research focused on the optimization of the receptive field and the use of color antagonism. However, these studies stayed at the level of the single visual cortex, only considering the feedforward effect of the visual pathway, which ignored the role of horizontal connection and feedback connection, making the models lack completeness. Based on these, this paper proposed a contour perception model that simulates the complex connection pattern of the visual cortex. First, the sparse coding characteristics of the LGN neurons were simulated to obtain the primary perception of the LGN on image contour. Second, the directional selectivity of the primary visual cortex receptive field was simulated, and the optimal orientation contour responses of the classical receptive field(CRF) were obtained. Third, simulating the hue perception characteristics of the advanced visual cortex, a three-channel hue perception model including surround suppression was constructed, which obtains the contour response through the action of the advanced visual cortex neurons. Finally, a visual cortex model fusing feedforward, feedback, and the horizontal connection was constructed, in which the horizontal connection part adjusted the enhancement and inhibition of neurons by simulating the windmill-like functional structure of the primary visual cortex. Using the interaction between three layers of neurons, we corrected the discharge strength of visual cortex neurons to improve the accuracy of contour extraction.

The contributions are as follows:

- (1) Simulated the windmill-like functional structure of the primary visual cortex, adjusting the enhancement and inhibition of the neuron action to simulate the horizontal connection of the visual cortex.
- (2) Simulated the advanced visual cortex hue perception characteristics to construct a three-channel hue perception model that includes surround suppression.
- (3) Constructed a contour perception model that integrates the feedforward, feedback, and horizontal connection patterns of the visual cortex.

The rest of this paper is organized as follows. In Section 2, we reviewed the related work in the past studies and summarized the advantages and limitations of the previous models. In Section 3, we presented the overall structure of the proposed model. In Section 4, we conducted the comparative test and analyzed the experiment results. In Sections 5 and 6, we presented the discussion and conclusion for the proposed model.

2 Related work

The innovations of the contour detection methods based on the biological vision mechanism can be roughly divided into three categories: optimization of the receptive field, application of the color antagonism, and connection mode of the visual cortex.

2.1 Optimization of the receptive field

Since Hubel and Wiesel [13] proposed the concept of the receptive field and pointed out that it has the characteristic of orientation selection, the receptive field has gradually become the focus of biologists and machine vision researchers. Based on the research of Hubel and Wiesel, Grigorescu [12] proposed to use the two-dimensional Gabor energy function to describe the orientation selection characteristics of the CRF, and further, used the difference of the Gaussian(DoG) function to simulate the lateral inhibition of the non-classical receptive field(nCRF). Since then, biological vision and mathematical models are linked together. In recent years, the researches on the form and function of the receptive field never stopped, Melotti [23] focused on the study of receptive field interaction and built a model to simulate the push-pull inhibition and surround inhibition of the receptive field. This model can suppress the texture while keeping the strongest responses to lines and edges. As for the researches on the form of the receptive field, Fang [11] inspired by the asymmetric receptive field of frogs, simulated the R3 cell model and proposed an information fusion strategy based on the bilateral asymmetric receptive fields. The R3 cell model can suppress the texture of the primary contour response of the image with different intensities of texture. Zhang [40] focused on the dynamic regulation of the receptive field under the external stimuli and the depth determination by binocular cells and proposed the DIDY model, which combined the luminance and disparity information responses based on the optimal orientation. Based on a determination of the relative importance of the receptive field responses, Liu [20] modeled the response of the stereoscopic image content of the classical and non-classical anisotropic receptive fields. However, the above contour detection models are unable to meet the needs of real scene image processing, which have some problems such as insufficient generalization ability and unsatisfactory extraction effect.

2.2 Application of the color antagonism

There are two complementary theories in the study of color vision, namely the theory of Trichromacy [28] and the theory of Opponent process [6]. They proposed the color antagonism theory, which has been applied to contour detection tasks by many researchers. Yang [34] proposed the SCO model based on the color antagonism mechanisms of the biological visual system. Compared with Yang's use of color antagonistic information only, Lin [17]

combined color information with brightness information and proposed a color antagonism model based on dark-light adaptivity, which used the dark and light channel structure to simulate the light and dark processing mechanism of the visual cortex. In addition, Yuan [37] simulated the human visual system and proposed a saliency model based on the color-opponent mechanism in the primary visual cortex. The researches on color antagonism only stayed at the level of the visual cortex, which can't reflect the complete process of image processing by the biological vision system.

2.3 Connection mode of the visual cortex

A complete biological vision model not only needs to simulate the internal mode in one visual cortex but also needs to consider the interaction between visual cortices. Due to the complexity of the biological vision pathway, its mathematical model is constantly updated with the progress of biological experiments. In recent studies, Capparelli [9] built a neural network in the primary visual cortex where two neural populations were organized in different layers within orientation hypercolumns that were connected by local, short-range, and long-range recurrent interactions. When trained with natural images, the model predicted a connectivity structure linking neurons with similar orientation preferences matching the typical patterns found for long-ranging horizontal axons and feedback projections in the visual cortex. Akbarinia [1] proposed a biologically-inspired edge detection model based on feedback and surround modulated, which modeled a contrast-dependent surround modulation of V1 receptive fields by accounting for full, far, iso-, and orthogonal-orientation surround. Cao [8] put forward a method for extracting local center-surround contrast information from nature images by using a normalized DoG function and a sigmoid activated function. Compared with previous contour detection models, this method can efficiently suppress textures more quickly and accurately. These models well simulated the structure of the biological visual pathways, but they all had the problems of large computational complexity, which led to the low computational efficiency.

3 Materials and methods

The biological vision system is a multi-channel, multi-cascade, and cross-channel complex information processing system [29]. For the sake of simplicity, most of the traditional calculation models paid attention to the construction of the neuron network structure within the functional areas, thus ignoring the connection mode between the layers. In reality, there are complex connection patterns within and between functional areas of the visual cortices [14], which are of great significance to the coding mechanism of the visual nervous system. Azzopardi [4] pointed out that the LGN is the center of the processing of visual information, which plays an important role in the initial processing and transmission of visual stimuli. Murgas [27] showed that the advanced visual cortex mainly plays a role in information processing and integration, which gathers the input signals from all parts. Therefore, focusing on the application of target contour perception, this paper simulated the information transmission relationship among the LGN, primary visual cortex, and advanced visual cortex in the visual pathway, and constructed a three-layer neural network to simulate the response and transmission process of the biological visual system. The contour perception model was shown in Fig. 1.

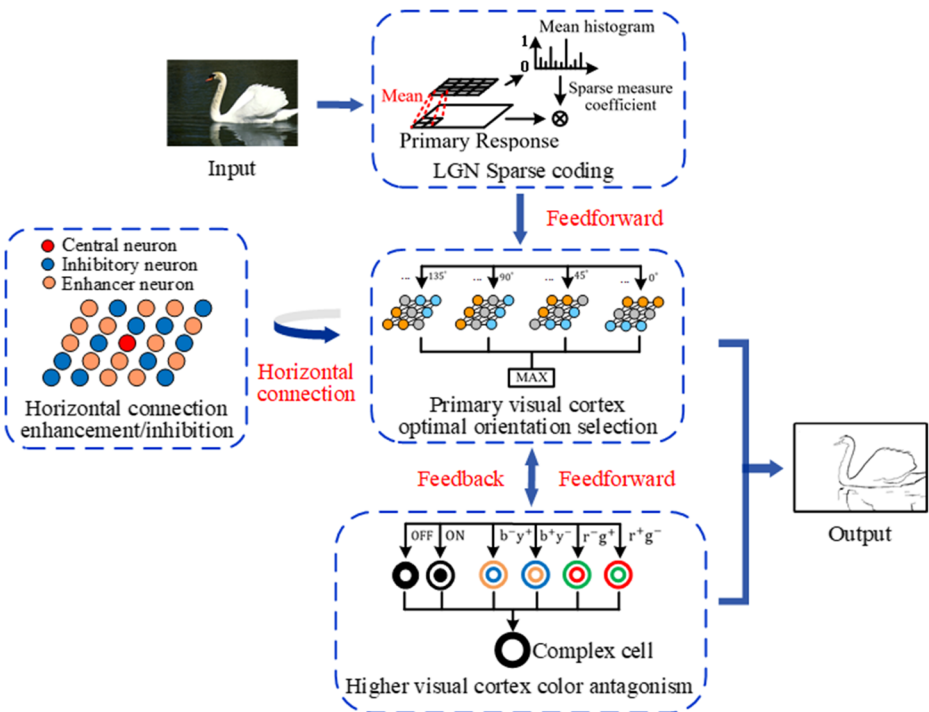


Fig. 1 A contour perception model that simulates the complex connection patterns of the visual cortex

3.1 LGN sparse measurement method based on statistical characteristics of the receptive field

LGN is the core area connecting the retina and the brain. The number of neurons in the LGN is significantly less than the anterior and posterior layers in the visual pathway, showing the sparse coding characteristics of visual information transmission. In the visual computing model, the common sparse measurement method has a limited effect on distinguishing the contour and texture regions. In order to optimize its distinguishing effect, this paper incorporated a weight factor into the sparse measurement simulated the neuron discharge coding of the LGN local area, and proposed a sparse measurement method based on the statistical characteristics of the receptive field.

In this paper, the sparse measurement method proposed by Alpert [2] is used, as shown in Formula (1) and (2). At the same time, defined the receptive field window S_{ij} , and used σ to represent the half-length of the receptive field window.

$$\tau \frac{d}{dt} E_{ij}^{LGN}(t) = -E_{ij}^{LGN}(t) + I_{ij} \tag{1}$$

$$spar_{ij} = \frac{1}{\sqrt{n} - 1} \left(\sqrt{n} - \frac{\|\vec{h}_{ij}\|_1}{\|\vec{h}_{ij}\|_2} \right) \tag{2}$$

In the formula, τ denotes the neuron membrane potential constant; E_{ij}^{LGN} denotes the membrane potential of neurons at the discharge coding model, (i, j) which simulates the LGN neuron; I_{ij} denotes the gray value of an external input image at (i, j) ; \vec{h}_{ij} denotes the histogram with E_{ij}^{LGN} in the receptive field window S_{ij} ; n denotes the dimension of \vec{h}_{ij} ; $\|\bullet\|_p$ denotes the p norm.

The sparse measurement method was proposed based on the image segmentation, which tests whether each region has texture features. Its purpose was to exclude the texture region from the segmentation point through the setting of sparsity and does not consider the extraction of the image contour. And so, the effect of applying it to contour optimization was limited. Therefore, this paper set the weight factor and sparse measurement threshold to improve the distinction between image contour and texture by neurons, as shown in Formulas (3–5).

$$\text{sparsity}_{ij} = \begin{cases} \frac{\delta_{ij}^2}{\mu_{ij}} \cdot \text{spar}_{ij} & , \text{ if } \text{spar}_{ij} \geq \text{threshold} \\ 0 & , \text{ if } \text{spar}_{ij} < \text{threshold} \end{cases} \quad (3)$$

$$\text{threshold} = \frac{\sum_i \sum_j \text{spar}_{ij}}{\sigma^2} \quad (4)$$

$$V_{ij}^{\text{LGN}} = E_{ij}^{\text{LGN}} \cdot \text{sparsity}_{ij} \quad (5)$$

In the formula, δ_{ij} , μ_{ij} denote the variance and mean of E_{ij}^{LGN} in the receptive field window S_{ij} . Variance denotes the inconsistency between pixels at (i, j) and surroundings - the greater the variance, the greater the probability that (i, j) is a contour. Mean denotes the brightness and darkness of the (i, j) -centered range of receptive fields. The smaller the mean is, the darker the image is. The variance item of the weight factor plays a role in distinguishing contour and background, and the mean item plays a role in balancing the brightness and darkness of the picture. *threshold* denotes the threshold of the sparse metric. Its value is the mean of sparse measurement in the receptive field window S_{ij} . V_{ij} denotes the neuronal membrane potential at (i, j) after sparse measurement.

3.2 Lateral adjustment of the V1 windmill-like receptive field

The visual cortex of higher mammals shows a clear orientation preference. Most neurons only respond to their preferred single orientation graphics. The functional columns of different preferences are arranged orderly in space. In the cortex, periodic expansion is carried out with a singular point as the center. This organizational structure of spatial arrangement according to cell function is called the windmill-like functional construction of the cortex [25]. Each subregion of the windmill is structurally differentiated and functionally divided [15] to adjust the input visual information. Most of the previous methods focused on the orientation selectivity and non-orientation selectivity inhibition of the primary visual cortex, these methods depended excessively on the judgment of the best orientation and were vulnerable to the regional noises. This paper argued that the enhancement and inhibition of neurons can be obtained by combining the windmill-like function construction of the cortex and by considering the optimal angle between two neurons in the receptive field, which can effectively avoid the error of the best orientation judgment caused by noises. Therefore, this paper constructed a lateral adjustment model

integrating the receptive field of the windmill structure to simulate the functional column characteristics of the windmill-like structure in the biological visual cortex.

The horizontal modulation of neurons at the windmill structure receptive field (i', j') on neurons at (i, j) is related to the distance between neurons and the angle between the optimal response direction of the neurons. Taking (i, j) as the center of the windmill, the angle of the optimal response direction of the two neurons was calculated. If the angle is in the range of $0 \sim \pi/4$ or $3\pi/4 \sim \pi$, it has an enhancement effect, and if the angle is in the range of $\pi/4 \sim 3\pi/4$, it has an inhibitory effect. The enhancement and inhibition of neurons in the receptive field at (i, j) are shown in Formula (6). Formula (7) and (8) represent the distance between neurons and the enhancement/inhibition function between neurons.

$$\begin{cases} \Delta E_{ij} = \sum_{(i', j') \in S_{ij}} D^{*} \omega_E \\ \Delta I_{ij} = \sum_{(i', j') \in S_{ij}} D^{*} \omega_I \end{cases} \tag{6}$$

$$D = \exp\left(-\frac{(i-i')^2 + (j-j')^2}{\sigma^2}\right) \tag{7}$$

$$\begin{cases} \omega_E = \exp\left(-\frac{|\phi_{ij} - \phi_{i'j'}|}{\sigma_{\Delta\phi}^2}\right), & 0 \leq |\phi_{ij} - \phi_{i'j'}| \leq \frac{\pi}{4} \\ \omega_E = \exp\left(-\frac{\pi - |\phi_{ij} - \phi_{i'j'}|}{\sigma_{\Delta\phi}^2}\right), & \frac{3\pi}{4} \leq |\phi_{ij} - \phi_{i'j'}| \leq \pi \\ \omega_I = \exp\left(-\frac{\frac{\pi}{2} - |\phi_{ij} - \phi_{i'j'}|}{\sigma_{\Delta\phi}^2}\right), & \frac{\pi}{4} \leq |\phi_{ij} - \phi_{i'j'}| \leq \frac{\pi}{2} \\ \omega_I = \exp\left(-\frac{|\phi_{ij} - \phi_{i'j'}| - \frac{\pi}{2}}{\sigma_{\Delta\phi}^2}\right), & \frac{\pi}{2} \leq |\phi_{ij} - \phi_{i'j'}| \leq \frac{3\pi}{4} \end{cases} \tag{8}$$

In the formulas, D denotes the distance between the neurons, the smaller the D is, the closer the distance between the two neurons is as well as the stronger the interaction is. $*$ denotes the convolution operation. ω_E, ω_I denotes the inter-neuronal enhancement and inhibition functions. $\sigma_{\Delta\phi}$ denotes the decay rate, which decides how fast the peripheral inhibition intensity decay as the orientation difference increases. φ_{ij} denotes the connection direction between (i, j) and (i', j') . $\varphi_{i'j'}$ denotes the best orientation of neurons at (i', j') , as shown in Fig. 2.

Primary visual cortex neurons have orientation selection characteristics and have different response intensities for different orientation stimulus signals. When the orientation of stimulus signals is equal to the optimal orientation of neurons, the responses of neurons are the strongest. In this paper, a two-dimensional Gabor function was used to simulate the orientation selection characteristics of CRF in the primary visual cortex, and the maximum response was selected as the edge response, as shown in Formula (9).

$$\begin{cases} g_{i,j}^{\sim}(\theta_k) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\tilde{i}^2 + \tilde{j}^2}{2\sigma^2}\right) \\ \theta_k = \frac{(k-1)\pi}{N_\theta}, \quad k = 1, 2, \dots, N_\theta \\ e_{ij}(\theta_k) = \left| \sum_{i', j' \in S_{ij}} I_{ij}(t) * \frac{\partial g_{i,j}^{\sim}(\theta_k)}{\partial i} \right| \\ V_{ij}^{\max} = \max(e_{ij}(\theta_k)) \end{cases} \tag{9}$$

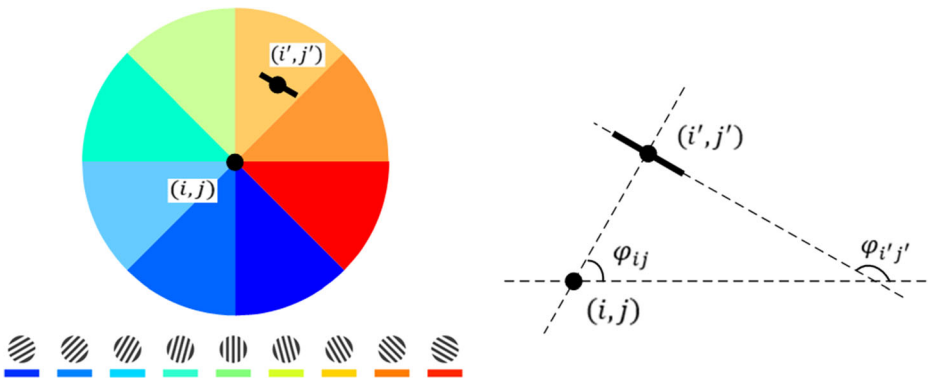


Fig. 2 The physical orientation between two neurons in connection and the optimal orientation angle of a certain neuron

In the formula, ε denotes the space compression ratio which controls the aspect ratio of the filter. $\tilde{i} = i \cos(\theta_k) + j \sin(\theta_k), \tilde{j} = -i \sin(\theta_k) + j \cos(\theta_k), \theta_k$ denotes the best orientation of the filter. k denotes the optimal orientation coefficient of the filter. N_θ denotes the number of orientations of the CRF filter. $e_{ij}(\theta_k)$ denotes the response of neurons (i, j) and the best orientation is θ_k . V_{ij}^{\max} denotes the neuronal membrane potential at (i, j) after the maximum response.

Although the main pathway of the biological visual system is perpendicular to the cortical surface, there are still highly convergent cluster pyramidal axons on the horizontal plane. When the target to be detected is located in a complex environment, the response intensity of the neurons caused by it depends on the modulation of the surrounding neurons. The lateral adjustment model of the integrated windmill-like structure receptive field proposed in this paper was used to achieve the effect of target contour enhancement and texture noises suppression by adjusting the horizontal connection between neurons, as shown in Formula (10).

$$V_{ij}^{HC} = V_{ij}^{\max} + \delta(\Delta E_{ij} - \Delta I_{ij}) \tag{10}$$

In the formula, V_{ij}^{\max} denotes the membrane potential of neurons at (i, j) after lateral regulation. δ denotes the intensity coefficient of neuronal interaction. $\Delta E_{ij}, \Delta I_{ij}$ denotes that the neurons at (i, j) was enhanced and inhibited in the receptive field, respectively.

3.3 Hue perception model of the advanced visual cortex

The advanced visual cortex receives the information from the primary visual cortex, which is fed back to the primary visual cortex after a series of actions, forming a circulatory regulation mechanism [26]. The traditional color antagonism model simply integrates the input of antagonistic cells based on the receptive field characteristics of neurons at all levels, without considering the odd and even channel structure of the visual pathway and the effect of complex cells on the input stimulation. Therefore, a three-channel advanced visual cortex hue perception system containing surround inhibition was constructed to simulate color antagonism and surround inhibition of complex cells in the advanced visual cortex, as shown in Fig. 3.

Considering the interaction between the CRF and nCRF in the visual cortex, the Gaussian difference function was integrated into the double antagonistic receptive field model. The

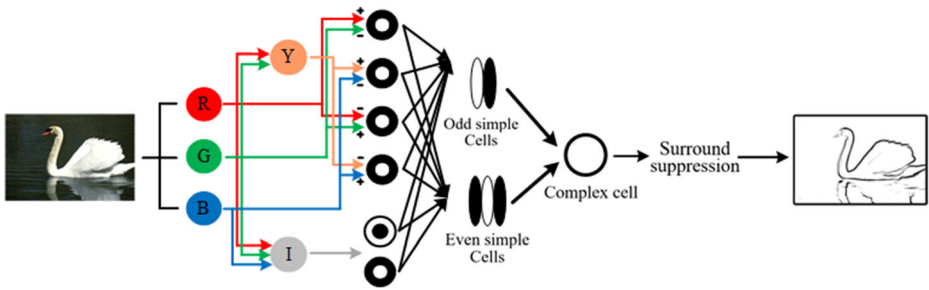


Fig. 3 Three-channel advanced visual cortex hue perception system with surround suppression

competition coefficient of neurons in the position (i, j) of r^+g^- channel is shown in Formula (11).

$$\begin{cases} DoG_{pq} = \frac{1}{\sqrt{2\pi}\sigma_E} \exp\left(-\frac{p^2+q^2}{2\sigma_E^2}\right) - \frac{1}{\sqrt{2\pi}\sigma_I} \exp\left(-\frac{p^2+q^2}{2\sigma_I^2}\right) \\ C_{ij}^{rg} = \frac{\sum_{p,q \in S_{ij}} [DoG_{pq}]^+ R_{pq} - \sum_{p,q \in S_{ij}} [-DoG_{pq}]^+ G_{pq}}{A_1 + \sum_{p,q \in S_{ij}} [DoG_{pq}]^+ R_{pq} + \sum_{p,q \in S_{ij}} [-DoG_{pq}]^+ G_{pq}} \end{cases} \quad (11)$$

In the formula, DoG_{pq} denotes the Gauss difference function of neurons at (p, q) . σ_E denotes the radius of the exciting receptive field. σ_I denotes the radius of the inhibitory receptive field. $[a]^+ = \max(0, a)$. R_{pq} , G_{pq} denotes the red and green component inputs of neurons at p and q , respectively. A_1 denotes the attenuation coefficient. For the g^+r^- channel, b^+y^- channel, and y^+b^- channel (i, j) position, the expression of neuronal activity is similar to Formula (11). Only by changing R_{pq} and G_{pq} as the corresponding color component inputs, were the neuronal competition coefficients C_{ij}^{gr} , C_{ij}^{by} and C_{ij}^{yb} at the (i, j) position of each channel obtained.

Since color information and brightness information play complementary roles in the distinction between contour and texture, the dark region is dominated by color information, and the bright region is dominated by brightness information. Therefore, the brightness channel is added on the basis of the color antagonistic channel to obtain the color information and brightness information at the same time. The brightness channel is divided into the opening channel and the closing channel. The opening channel is responsible for enhancing the information of brightness higher than the surrounding area, setting $\sigma_E < \sigma_I$. The closing channel is responsible for suppressing the information of brightness lower than the surrounding area, setting $\sigma_E > \sigma_I$. Letting $R_{ij} = G_{ij} = I_{ij}$, Formula (11) can obtain the neuron competition coefficients C_{ij}^{on} and C_{ij}^{off} at the (i, j) position of the on and off paths, respectively.

The feature extraction of a visual system is a calculation process of local energy. If the activity of three-channel neurons were simply added, it will cause confusion between the contour and detail area. The traditional local energy model uses the square root of the square sum of the response of the Gabor filter with two orthogonal phases to represent the Gabor energy and directly fuses the output of odd and even filters. There is no advantage in image detail processing. Inspired by the multi-channel filtering theory for the early processing of visual information in the human visual system [24], this paper proposes to simulate the structure of odd and even channels in the visual cortex, and applies the local energy model of multi-channel filtering to integrate the obtained color antagonism and brightness antagonism information.

The two-dimensional Gabor filter was used to filter the input information, as shown in Formula (12). The information obtained by the filter was combined with the neuronal activity of each channel to obtain the simple cell activity of each odd and even component. g^{+r} channel (i,j) position's simple cell activity of odd and even components were shown in Formula (13).

$$g_{ij} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{i^2 + \varepsilon^2 j^2}{2\sigma^2}\right) \cos\left(2\pi \frac{i}{\lambda} + \phi\right) \tag{12}$$

$$\begin{cases} o_{ij}^{rg} = \frac{\sum_{p,q \in S_{ij}} g_{pq}^{O,rg} c_{pq}^{rg}}{A_2 + \sum_{p,q \in S_{ij}} |g_{pq}^{O,rg}| c_{pq}^{rg}} \\ e_{ij}^{rg} = \frac{\sum_{p,q \in S_{ij}} g_{pq}^{E,rg} c_{pq}^{rg}}{A_2 + \sum_{p,q \in S_{ij}} |g_{pq}^{E,rg}| c_{pq}^{rg}} \end{cases} \tag{13}$$

In the formula, φ denotes phase parameters, for odd symmetric filter $\varphi = -\pi/2$ or $\pi/2$ and for even symmetric filter $\varphi = 0$ or π . $g^O pq$, $g^E pq$ denotes the odd/even components at the (p,q) position, respectively. For the g^{+r} channel, b^{+y} channel, and y^{+b} channel, the simple cell activity expression of odd and even components at position (i,j) were similar to Formula (13). One only needed to modify C_{ij}^{rg} to C_{ij}^{gr} , C_{ij}^{by} and C_{ij}^{yb} to obtain the odd and even components of each channel o_{ij}^{gr} , e_{ij}^{gr} , o_{ij}^{by} , e_{ij}^{by} and o_{ij}^{yb} , e_{ij}^{yb} . For the simple cell activity of odd and even components of the brightness channel, it was necessary to fuse the open channel and the closed channel, as shown in Formula (14).

$$\begin{cases} o_{ij}^L = \frac{\sum_{p,q \in S_{ij}} g_{pq}^{O,LO} ([c_{ij}^{on}]^+ - [c_{ij}^{off}]^+)}{A_2 + \sum_{p,q \in S_{ij}} |g_{pq}^{O,LO}| ([c_{ij}^{on}]^+ + [c_{ij}^{off}]^+)} \\ e_{ij}^L = \frac{\sum_{p,q \in S_{ij}} g_{pq}^{E,LO} ([c_{ij}^{on}]^+ - [c_{ij}^{off}]^+)}{A_2 + \sum_{p,q \in S_{ij}} |g_{pq}^{E,LO}| ([c_{ij}^{on}]^+ + [c_{ij}^{off}]^+)} \end{cases} \tag{14}$$

Only considering the role of simple cells may lead to incomplete contour extraction and excessive texture in the details. Therefore, this paper used two complex layers of cells to simulate the role of the advanced visual cortex and process the input information from simple cells. The first layer was responsible for fusing ten groups of simple cell responses of three channels and unifying the contour features extracted from each channel. The second layer adopted the surround suppression and used the competitive network in the surrounding environment outside the cell receptive field to achieve the texture suppression effect, as shown in Formula (15) and (16).

$$v_{ij} = \sum_{k=L,rg,gr,by,yb} \left([o_{ij}^k]^+ + [e_{ij}^k]^+ \right) \tag{15}$$

$$V_{ij}^G = \frac{v_{ij} - \varsigma \cdot \sum_{p,q \in S_{ij}} (\text{DoG}_{pq} \cdot v_{pq})}{A_3 + v_{ij} + \sum_{p,q \in S_{ij}} (\text{DoG}_{pq} \cdot v_{pq})} \tag{16}$$

In the formulas, V_{ij}^G denotes the neuronal membrane potential at (i,j) after hue perception. A_3 denotes the model coefficient. ς denotes inhibition constant.

3.4 Contour extraction model that simulates the complex connection of visual pathway

Physiological studies have shown that the complex and changeable pathway structure is the basis for the formation of the visual system. Most traditional visual nerve calculation models focused on the construction of the neural network within functional areas, ignoring the interaction between functional areas. In order to better simulate the biological vision system, this paper proposed a contour perception model to simulate the complex connection pattern of the visual cortex. A three-layer neural network was constructed to simulate the response and transmission process of LGN, the primary visual cortex, and the advanced visual cortex. The output was integrated with the response of advanced visual cortex to correct background contour and texture noises, as shown in Formula (17).

$$E_{ij} = A \cdot \left[e^{-t/\tau} + \left(\frac{\alpha \cdot V_{ij}^{LGN} + \beta \cdot V_{ij}^G + \gamma \cdot V_{ij}^{HC}}{\tau} \right) t \right] + B \cdot V_{ij}^G \tag{17}$$

In the formula, E_{ij} denotes the output contour response. α , β and γ denotes the feedforward, feedback, and horizontal connection coefficient, respectively. A , B denote the coefficient of integration, respectively.

3.5 Parameters description

As this paper contained many parameters, we added Table 1 to illustrate the definitions and values of related parameters used in this paper.

4 Experiments

In order to verify the effectiveness of the proposed method, the experiments selected the mainstream biological vision methods and the deep learning methods for comparison. The

Table 1 Definitions and values of related parameters used in this paper

Parameter	Parameter definitions	Value
τ	Neuron membrane potential constant	70
σ	Size of the receptive field of primary visual cortex neurons	1.5
$\sigma_{\Delta\phi}$	Decay rate	1.5
N_θ	Number of orientations of the CRF filter	8
ε	Ellipticity of the filter	0.5
δ	Intensity coefficient of neuronal interaction	0.5
A_1, A_2, A_3	Model coefficient	1
ς	Inhibition constant	0.5
α	Feedforward connection coefficient	0.5
β	Feedback connection coefficient	0.25
γ	Horizontal connection coefficient	0.25

biological vision methods included the contour detection model based on the multiple-cue inhibition (MCI) [33], the contour detection model based on DO and spatial sparseness constraint (SCO) [34], the boundary detection model based on the feedback and surround modulation (SED) [1], and the contour detection model based on the bilateral asymmetric receptive field (BAR) [11]. As for deep learning methods, including the HED [32] network and the RCF [18] network. The BSDS500 [3] dataset and the NYUD dataset were selected for comparative analysis of each method.

4.1 BSDS500 dataset

The BSDS500 dataset contains 500 images. The corresponding 500 ground-truth maps are the contour of the artificial identification, which are used to evaluate the effectiveness of the method.

4.1.1 Coefficient selection experiments

In order to determine the influence of different values of the fusion coefficients A and B in Eq. 17 on the overall performance of the model, the coefficient selection experiment was carried out. The accuracy rate, recall rate and the F-score evaluation index proposed in the literature [22] were used to quantitatively analyze the results. The specific calculation process of the F-score is shown in Formulas (18–20). Since there is a deviation between the final contour map and the standard contour map, this paper defined that if the detected pixel appears in the 5×5 neighborhood of the standard contour pixel, it is judged that the pixel is labeled correctly. The standard contour pixel set is denoted by E_D , the contour set detected by the algorithm is denoted by E_{GT} . The correct pixel set detected by the algorithm is $E = E_D \cap (E_{GT} \oplus T)$, where \oplus denotes the expansion operation and T denotes the 5×5 structural unit. The error pixel set E_{FP} detected by the algorithm is $E_{FP} = E_D - E$. The missing pixel set E_{FN} of the algorithm is $E_{FN} = E_{GT} - (E_{GT} \cap (E_D \oplus T))$.

$$P_r = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FP})} \quad (18)$$

$$R_c = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FN})} \quad (19)$$

$$F = \frac{2P_r R_c}{P_r + R_c} \quad (20)$$

In the formula, $\text{card}(E)$ denotes the number of elements in the set E , P_r denotes the accuracy, and R_c denotes the recall rate. F denotes the consistency between the contour detected by the algorithm and the standard contour. A series of accuracy and recall rates can be obtained by adjusting the threshold. With the recall rate as the horizontal axis and the accuracy as the vertical axis, the precision–recall(P-R) curve is plotted, which is shown in Fig. 4. The larger the area under the P-R curve, the better the contour extraction performance of the algorithm.

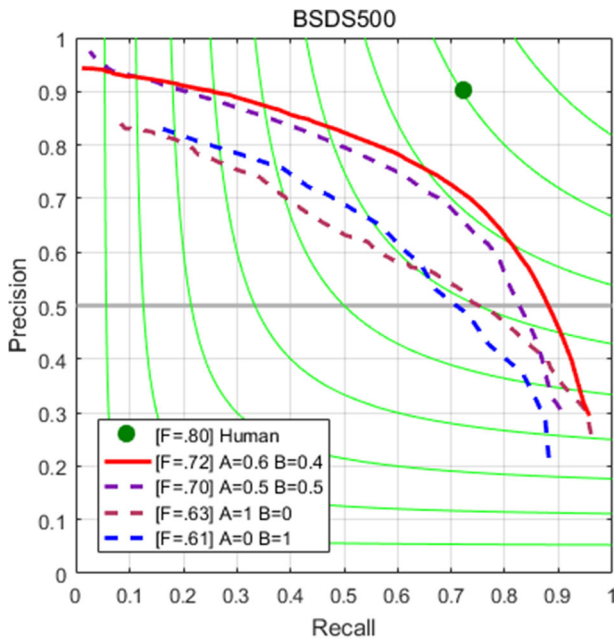


Fig. 4 P-R curves of different coefficients for the BSDS500 dataset. The notion of [F=] in the legend illustrates the best F-score

In the case of determining the fusion coefficients A and B , six images were randomly selected from the BSDS500 dataset as the experimental objects. It can be seen from the figure and data that when $A = 0.6$ and $B = 0.4$, the proposed method has a better detection effect.

4.1.2 Comparative experiments

In order to verify the effectiveness of the proposed method in this paper, the BSDS500 dataset was used for comparative experiments, the results are shown in Fig. 5. In order to show the contour extraction effect of each method more clearly, the optimal binarization operation was carried out on each contour map. We named the method proposed in this paper as CCP.

It can be seen from the results in Fig. 5 that the MCI method combined a variety of visual features and can better extract the key contour information, but the processing of the details was insufficient, such as the 8068 red circle area. SCO method extracted contours based on the color information and added the texture suppression part, which can better balance the contour extraction and the textures suppression. However, in some complex images, the contours and the textures still cannot be completely separated, such as the 176,035 red circle area. On the basis of the MCI method, the BAR method added the asymmetric receptive field mechanism, so its detection effect was slightly improved, but still had many textures, such as the 113,044 red circle area. The contour obtained by the SED method was relatively complete, but there are some textures around the contour, such as the 38,092 and 118,035 red circle areas. CCP method took the feedforward, feedback, and horizontal connection into consideration, using visual cortex multi-level interaction and dynamic connection to finish the contour extraction task. So the obtained contour was more complete than other methods. Furthermore, there were relatively fewer textures.

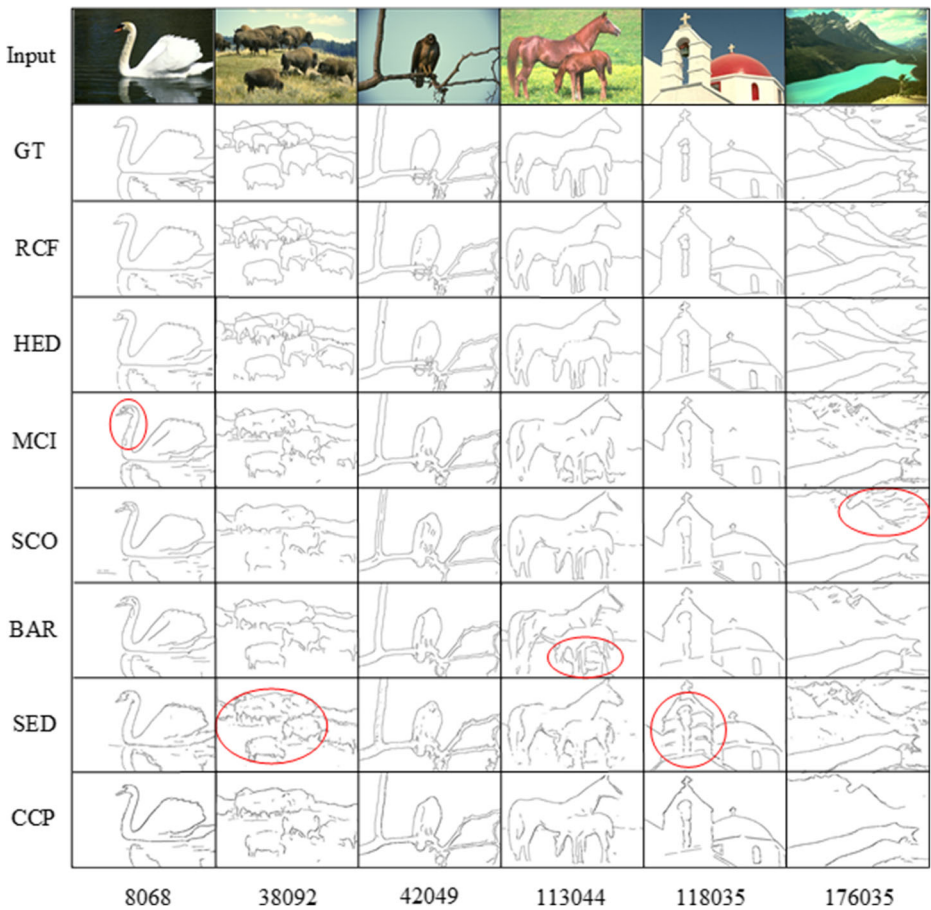


Fig. 5 Contour detection results of the BSDS500 dataset

For the contour obtained by each method in the BSDS500 dataset, uses accuracy rate and recall rate to quantitatively analyze the results. Figure 6 shows the P-R curves and F-scores of different algorithms. Table 2 is the performance evaluation of each method applied to the BSDS500 dataset-optimal dataset scale (ODS), optimal image scale (OIS), and the average accuracy (AP) [3] of each method.

It can be seen from the graph and table that the CCP method performs better than the methods based on the biological vision (MCI, SCO, BAR and SED) in various performance indicators. As for deep learning methods such as RCF and HED, the CCP method is slightly insufficient in performance indicators, but deep learning methods have many limitations. First of all, the results obtained through deep learning methods cannot be corrected, and the portability is weak, if the dataset is replaced, it needs to be retrained from scratch. Secondly, it is generally accepted that the deep learning methods such as RCF and HED are more of a network structure at the black box level, lacking the interpretability of brain-like methods. To a certain extent, deep learning methods also lose the opportunity for humans to find errors. Finally, due to the complexity of the models in deep learning, the time complexity of the algorithms has increased dramatically. In order to ensure the real-time performance of the

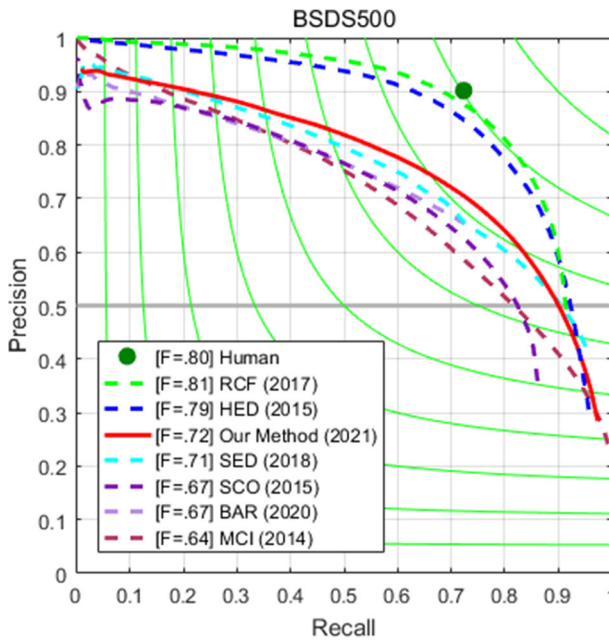


Fig. 6 P-R curves of different models for the BSDS500 dataset

algorithm, higher programming skills, more training time and better hardware support are required. In contrast, the method proposed in this paper used a visual neural computing model to simulate the mechanism of visual information transmission and processing in the visual pathway. Therefore, CCP method had better feasibility in the smaller dataset scale.

4.1.3 Ablation experiments

In order to verify the contribution of each part of the method to contour extraction, ablation experiments were carried out. The complete model was compared with the removal of the horizontal connection module, the removal of the feedback connection module, and the removal of the horizontal connection and feedback connection module at the same time. The P-R curve is shown in Fig. 7, and the indexes of ODS, OIS, and AP are shown in Table 3.

Table 2 Comparison of the models for the BSDS500 dataset

Method	ODS	OIS	AP
Human	0.8	0.8	—
Deep Learning			
HED(2015)	0.78	0.80	0.83
RCF(2019)	0.81	0.82	0.84
Biological			
MCI(2014)	0.64	0.66	0.56
SCO(2015)	0.67	0.71	0.69
SED(2018)	0.71	0.73	0.71
BAR(2020)	0.67	0.69	0.66
CCP(2022)	0.72	0.74	0.71

The bold indicates the optimal indicator

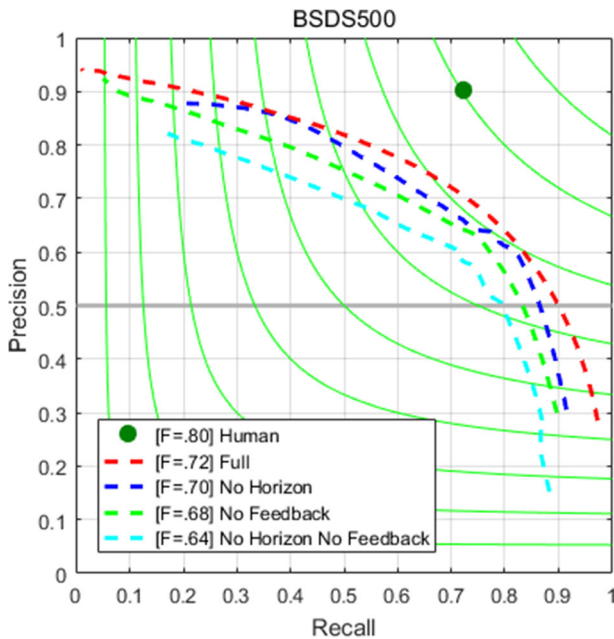


Fig. 7 P-R curves of ablation experiments for the BSDS500 dataset

It can be seen from the curve in Fig. 7 and the data in Table 3 that both the horizontal connection module and the feedback connection module contribute to the model constructed in this paper. When the horizontal connection module or the feedback connection module is removed, the F-score drops to 0.70 and 0.68, respectively. When the horizontal connection and feedback connection module are removed at the same time, the F-score drops to 0.64. The horizontal connection module simulates the lateral adjustment function of the biological visual receptive field, which helps to avoid the influence of regional noise and enhance the robustness of the model. The feedback connection module simulates the feedback process of the biological visual system, which uses the information obtained by the advanced visual cortex to adjust the primary visual cortex to enhance the overall stability of the model.

4.2 NYUD dataset

In order to further demonstrate the effectiveness of the model algorithm in this paper, the NYUD dataset was used for further analysis. The NYUD dataset consists of video sequences of various indoor scenes recorded by Microsoft Kinect's RGB and depth cameras, which contain 1449 annotated RGB images and depth maps from 464 scenes in 3 cities. In this paper,

Table 3 Results of ablation experiments for the BSDS500 dataset

Method	ODS	OIS	AP
Full	0.72	0.74	0.71
No Horizon	0.70	0.72	0.68
No Feedback	0.68	0.69	0.67
No Horizon No Feedback	0.64	0.65	0.60

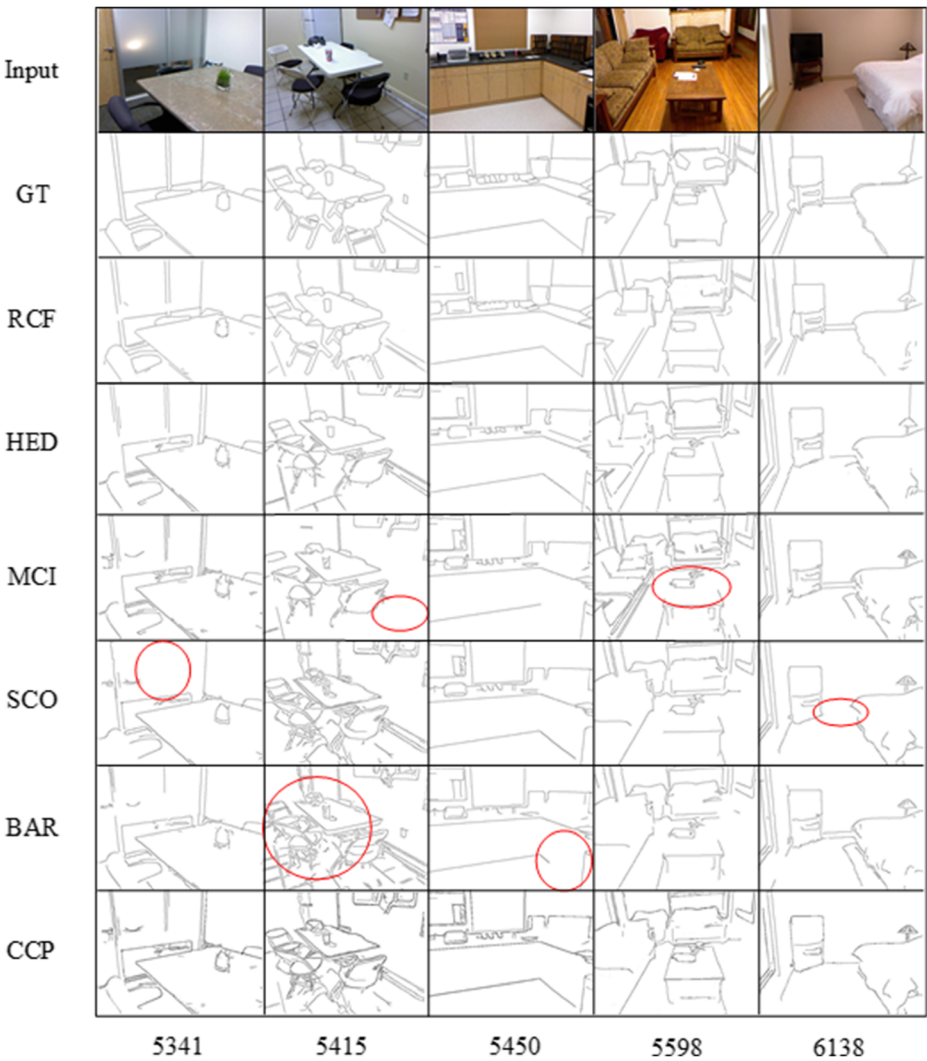


Fig. 8 Contour detection results for the NYUD dataset

five pictures were randomly selected from the NYUD dataset for results display and data analysis, as shown in Fig. 8.

It can be seen from the results in Fig. 8, the MCI method can extract the contour of the image relatively complete, but for areas with weak contrast characteristics, such as the red frame area of 5415 and 5598, the contours extracted by the MCI method were missing. SCO method used color information to extract the contour, it was easy to ignore the contour information where the color difference was small, such as the red frame area of 5341 and 6138. BAR method cannot balance the contour extraction and the texture suppression, for example, the red frame area of 5415 has more texture information, while the red frame area of 5450 has obvious missing the contours. Due to the richness of the image details in the NYUD dataset, the contours obtained by the CCP method cannot satisfy higher completeness and accuracy, but compared with the three biological vision-based

methods of MCI, SCO and BAR, the contours obtained by the CCP method were more complete, as well as the texture information is relatively less.

Using the method proposed by Grigorescu [12] to quantitatively evaluate the obtained contour. The performance evaluation index error rate e_{FP} , missed rate e_{FN} , and overall performance index Performance-value were calculated by Formulas (21–23). The specific results are shown in Table 4.

$$e_{FP} = \frac{\text{card}(E_{FP})}{\text{card}(E)} \quad (21)$$

$$e_{FN} = \frac{\text{card}(E_{FN})}{\text{card}(E_{GT})} \quad (22)$$

$$\text{Performance} = \frac{\text{card}(E)}{\text{card}(E) + \text{card}(E_{FP}) + \text{card}(E_{FN})} \quad (23)$$

For the more complex image set of NYUD, the MCI method can extract image contour information well, but the false detection rate e_{FP} is higher on the images with more details. The Performance-value of the BAR method is slightly higher than that of the MCI method, but its contour extraction is not complete. SCO method used the color information and had good performance for most images in the NYUD dataset, but it was difficult to balance the contour extraction and the texture suppression for images with dim brightness and a large number of details. In this paper, the overall performance index Performance-value was improved compared with other methods, meanwhile, the false detection rate and missed detection rate were decreased. According to the above analysis, the CCP method performs better both in contour extraction and texture suppression.

By using each method to extract the contours of the images in the NYUD dataset, 1449 best Performance-values can be obtained. In order to verify the stability of each method, the Performance-values obtained through HED, RCF, MCI, SCO, BAR and CCP were statistically calculated in the form of box-and-whisker statistics. The top and bottom of the box represent the maximum and minimum values after removing the outliers, and the middle horizontal line represents the median of the Performance-value. The shorter the box, the better the stability of the method [36] (Fig. 9).

Table 4 Comparison of the models for the NYUD dataset

Method	e_{FP}	e_{FN}	Performance
Deep Learning			
HED(2015)	0.84	0.36	0.47
RCF(2019)	0.81	0.34	0.49
Biological			
MCI(2014)	1.24	0.41	0.35
SCO(2015)	0.93	0.48	0.36
BAR(2020)	1.12	0.44	0.36
Our Method	0.91	0.39	0.42

The bold indicates the optimal indicator

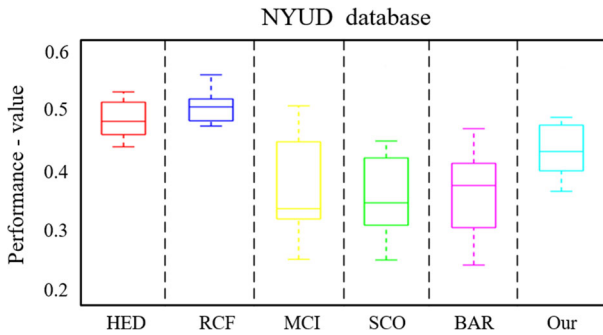


Fig. 9 Box-and-Whisker statistics of Performance-value in contour detection for the NYUD dataset

It can be seen from the box-and-whisker statistics that the deep learning methods (HED, RCF) are superior in performance to the methods based on the biological vision. However, the acquisition of its excellent performance required a large amount of training time. If the dataset is replaced, they need to be retrained from scratch. On the contrary, the methods based on biological vision do not require training, and the models are constructed by simulating the biological vision system, which are fully interpretable. When comparing the methods based on the biological vision, it can be found that the CCP method gets a higher Performance-value, which means that the CCP method has good performance indicators. Also, the box obtained by the paper's method is shorter than others, which means that the method is more stable.

5 Discussion

In this paper, we chose the BSDS500 dataset and NYUD dataset to compare the advantages and limitations of the proposed method, mainstream biological vision methods and popular deep learning methods.

As for the methods based on biological vision, the MCI method fully considers the influence of various factors on contour extraction. However, it excessively relies on the local features of the image, so the effect of contour extraction on texture boundaries and low-contrast regions needs to be improved. The SCO method constructed a contour detection model based on color antagonism and spatial sparsity constraint suppression. Nevertheless, it can only reflect the local contrast of the images and ignore the overall characteristics. Based on the feedback and surround modulated, the SED method constructed a contour detection model which structured a contrast-dependent surround modulation of V1 receptive fields. The contours obtained by this method are still having numerous textures at the boundary. The BAR method constructed a contour detection model based on the bilateral asymmetric receptive fields of the visual pathway, helping to highlight the contrast difference in local areas. However, this method only considers the brightness and contrast information, ignoring the rest of the features, which makes it unable to balance the contour extraction and texture suppression.

The CCP method simulated the complex connection pattern of the visual cortex, taking feedforward, feedback, and horizontal connection into consideration, to construct a contour calculation model inspired by the biological vision. In addition, the application of sparse coding, optimal orientation selection and other visual mechanisms helped the model to obtain the full contours efficiently. Furthermore, benefitting from the use of hue perception based on

the surround suppression mechanism, the CCP method produced a better suppression effect on textures. But there are still some limitations, the CCP method was essentially an unsupervised learning pattern, which didn't require labeling of the sample set, so it lacked the necessary prior knowledge or experience for the data. Therefore, after using the difference in the receptive field response to extract the primary contour, the subsequent processing cannot improve the closure and continuity of the target contour. In addition, the CCP method did not consider the binocular disparity structure. When processing images containing depth information, some important information will be lost, resulting in incomplete contour lines.

6 Conclusion

Based on the biological vision, this paper discussed the influence of the interaction of different visual cortexes on the visual information transmission. First, the sparse coding characteristic of the LGN was simulated to obtain the primary perception of LGN neurons on the contour. Second, the directional selectivity of CRF in the primary visual cortex was simulated, the lateral adjustment model integrating windmill-like structure receptive fields was constructed to obtain the perception of the primary visual cortex on image contour. Third, by simulating the hue perception characteristics of the advanced visual cortex, a three-channel hue perception model including surround suppression was constructed. Through the fusion of the complex cells' response to color antagonism channel and brightness channel and its surround suppression characteristics, we obtained the response of the advanced visual cortex to image contour, which effectively highlights the target contour as well as suppresses texture noise. Finally, a contour detection model was constructed to simulate the complex connection of visual cortex. The connection includes the feedforward input from LGN, the feedback input from the advanced visual cortex and the horizontal input from the same layer of neurons. The interaction between the visual cortexes was used to correct the discharge strength of the visual cortex neurons for input image stimulus and improves the accuracy of the contour.

Some improvements to the CCP model can be implemented in future work. (1) Developing the role of labeled samples in the contour recognition process, and exploring the construction and implementation of the semi-supervised learning methods for the visual computing models. In order to realize the image contour extraction based on the certain prior knowledge. (2) Introduce the binocular disparity structure into the visual computing model to extract the depth information of the target images, so as to achieve the accurate and complete expression of contours.

Acknowledgements This work has been supported by the Laboratory of Pattern Recognition and Image Processing in Hangzhou Dianzi University.

Declarations We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled "A contour perception model that simulates the complex connection pattern of the visual cortex".

References

1. Akbarinia A, Parraga CA (2018) Feedback and surround modulated boundary detection[J]. *Int J Comput Vision* 126(12):1367–1380
2. Alpert S, Galun M, Brandt A et al (2011) Image segmentation by probabilistic bottom-up aggregation and cue integration[J]. *IEEE Trans Pattern Anal Mach Intell* 34(2):315–327
3. Arbelaez P, Maire M, Fowlkes C et al (2010) Contour detection and hierarchical image segmentation[J]. *IEEE Trans Pattern Anal Mach Intell* 33(5):898–916
4. Azzopardi G, Petkov N (2012) A CORF computational model of a simple cell that relies on LGN input outperforms the Gabor function model[J]. *Biol Cybern* 106(3):177–189
5. Bertasius G, Shi J, Torresani L (2015) Deepedge: a multi-scale bifurcated deep network for top-down contour detection[C]. *Proceedings of the IEEE conference on computer vision and pattern recognition*. IEEE, Boston, pp 4380–4389
6. Buchsbaum G, Gottschalk A (1983) Trichromacy, opponent colours coding and optimum colour information transmission in the retina[J]. *Proc R Soc Lond Ser B Biol Sci* 220(1218):89–113
7. Canny J (1986) A computational approach to edge detection[J]. *IEEE Trans Pattern Anal Mach Intell* 8(6):679–698
8. Cao YJ, Lin C, Pan YJ et al (2019) Application of the center-surround mechanism to contour detection[J]. *Multimed Tools Appl* 78(17):25121–25141
9. Capparelli F, Pawelzik K, Ernst U (2019) Constrained inference in sparse coding reproduces contextual effects and predicts laminar neural dynamics[J]. *PLoS Comput Biol* 15(10):e1007370
10. Dollár P, Zitnick CL (2013) Structured forests for fast edge detection[C]. *Proceedings of the IEEE international conference on computer vision*. IEEE, Sydney, pp 1841–1848
11. Fang T, Fan Y, Wu W (2020) Salient contour detection on the basis of the mechanism of bilateral asymmetric receptive fields[J]. *SIViP* 14(7):1461–1469
12. Grigorescu C, Petkov N, Westenberg MA (2003) Contour detection based on nonclassical receptive field inhibition[J]. *IEEE Trans Image Process* 12(7):729–739
13. Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex[J]. *J Physiol* 160(1):106–154
14. Jacob T, Snyder W, Feng J et al (2016) A neural model for straight line detection in the human visual cortex[J]. *Neurocomputing* 199:185–196
15. Li M, Song XM, Xu T et al (2019) Subdomains within orientation columns of primary visual cortex[J]. *Sci Adv* 5(6):eaaw0807
16. Lim JJ, Zitnick CL, Dollár P (2013) Sketch tokens: a learned mid-level representation for contour and object detection[C]. *Proceedings of the IEEE conference on computer vision and pattern recognition*. IEEE, Portland, pp 3158–3165
17. Lin C, Zhao HJ, Cao YJ (2019) Improved color opponent contour detection model based on dark and light adaptation[J]. *Autom Control Comput Sci* 53(6):560–571
18. Liu Y, Cheng MM, Hu X et al (2017) Richer convolutional features for edge detection[C]. *Proceedings of the IEEE conference on computer vision and pattern recognition*. IEEE, Hawaii, pp 3000–3009
19. Liu Z, Zhang W, Zhao P (2020) A cross-modal adaptive gated fusion generative adversarial network for RGB-D salient object detection[J]. *Neurocomputing* 387:210–220
20. Liu L, Zhang J, Saad MA et al (2020) Blind S3D image quality prediction using classical and non-classical receptive field models[J]. *Sig Process Image Commun* 87:115915
21. Ma W, Gong C, Xu S et al (2020) Multi-scale spatial context-based semantic edge detection[J]. *Inf Fusion* 64:238–251
22. Martin DR, Fowlkes CC, Malik J (2004) Learning to detect natural image boundaries using local brightness, color, and texture cues[J]. *IEEE Trans Pattern Anal Mach Intell* 26(5):530–549
23. Melotti D, Heimbach K, Rodríguez-Sánchez A et al (2020) A robust contour detection operator with combined push-pull inhibition and surround suppression[J]. *Inf Sci* 524:229–240
24. Mingolla E, Ross W, Grossberg S (1999) A neural network for enhancing boundaries and surfaces in synthetic aperture radar images[J]. *Neural Netw* 12(3):499–511
25. Mohan YS, Jayakumar J, Lloyd EKJ et al (2019) Diversity of feature selectivity in macaque visual cortex arising from a limited number of broadly tuned input channels[J]. *Cereb Cortex* 29(12):5255–5268
26. Moratti S, Méndez-Bértolo C, Del-Pozo F et al (2014) Dynamic gamma frequency feedback coupling between higher and lower order visual cortices underlies perceptual completion in humans[J]. *NeuroImage* 86:470–479
27. Murgas KA, Wilson AM, Michael V et al (2020) Unique spatial integration in mouse primary visual cortex and higher visual areas[J]. *J Neurosci* 40(9):1862–1873
28. Nathans J, Thomas D, Hogness DS (1986) Molecular genetics of human color vision: the genes encoding blue, green, and red pigments[J]. *Science* 232(4747):193–202

29. Palmerston JB, Zhou Y, Chan RHM (2020) Comparing biological and artificial vision systems: network measures of functional connectivity[J]. *Neurosci Lett* 739:135407
30. Wang Y, Zhao X, Li Y et al (2018) Deep crisp boundaries: From boundaries to higher-level tasks[J]. *IEEE Trans Image Process* 28(3):1285–1298
31. Wu R, Feng M, Guan W et al (2019) A mutual learning method for salient object detection with intertwined multi-supervision[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Long Beach, pp 8150–8159
32. Xie S, Tu Z (2015) Holistically-nested edge detection[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, Boston, pp 1395–1403
33. Yang KF, Li CY, Li YJ (2014) Multifeature-based surround inhibition improves contour detection in natural images[J]. *IEEE Trans Image Process* 23(12):5020–5032
34. Yang KF, Gao SB, Guo CF et al (2015) Boundary detection using double-opponency and spatial sparseness constraint[J]. *IEEE Trans Image Process* 24(8):2565–2578
35. Yang H, Li Y, Yan X et al (2019) ContourGAN: Image contour detection with generative adversarial network[J]. *Knowl Based Syst* 164:21–28
36. Yoav B (1988) Opening the Box of a Boxplot. *Am Stat* 42(4):257. <https://doi.org/10.2307/2685133>
37. Yuan B, Han L, Yan H (2021) Explore double-opponency and skin color for saliency detection[J]. *Neurocomputing* 425:219–230
38. Zhang H, Sindagi V, Patel VM (2019) Image de-raining using a conditional generative adversarial network[J]. *IEEE Trans Circuits Syst Video Technol* 30(11):3943–3956
39. Zhang Y, Tian Y, Kong Y et al (2020) Residual dense network for image restoration[J]. *IEEE Trans Pattern Anal Mach Intell* 43(7):2480–2495
40. Zhang Q, Lin C, Li F (2021) Application of binocular disparity and receptive field dynamics: A biologically-inspired model for contour detection[J]. *Pattern Recog* 110:107657

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.