Edge Detection Based On the Characteristic of Primary Visual Cortex Cells

M M Zhu¹*, Y L Xu¹ and H Q Ma¹

¹Air Force Engineering University, Xi’an, China

*ming_paper@163.com

Abstract. Aiming at the problem that it is difficult to balance the accuracy of edge detection and anti-noise performance, and referring to the dynamic and static perceptions of the primary visual cortex (V1) cells, a V1 cell model is established to perform edge detection. A spatiotemporal filter is adopted to simulate the receptive field of V1 simple cells, the model V1 cell is obtained after integrating the responses of simple cells by half-wave rectification and normalization, Then the natural image edge is detected by using static perception of V1 cells. The simulation results show that, the V1 model can basically fit the biological data and has the universality of biology. What’s more, compared with other edge detection operators, the proposed model is more effective and has better robustness.

1. Introduction

Biological vision system in the image processing has the superiority that the current computer vision system can’t match. Therefore, it is an inevitable trend to use the research results of biological vision to promote the development of computer vision. With the development of brain science and physiological science, scholars have been studying the mechanism of biological vision system, which is applied to the research of image retrieval and image recognition, and has achieved many practical applications [1]. In the field of image processing, edge detection as one of the most basic technology has an important role in the subsequent image processing process. At present, the classical edge detection algorithm with image gradient or second derivative uses different differential operators to get the edge of image, such as Sobel operator, Log operator, Canny operator [2]. Some new edge detection methods have been proposed, such as wavelet multi-scale, fractal theory, mathematical morphology and so on, as scholars continue to explore in the field of image edge detection. However, most of the existing algorithms have a question about detection accuracy and anti-noise performance balance. Although there are some ways to solve this problem, those methods greatly increase the amount of computation. Therefore, finding an edge detection algorithm, which holds simple algorithm, good detection performance and strong anti-noise performance, has been an important research direction of image processing [3].

A large number of studies have shown that the biological vision system consists of dorsal pathway and ventral pathway whose structures and functions are relatively independent [4]. As a common starting point for both pathways, V1 has both dynamic and static feature-aware capabilities. Hubel and Wiesel [5] pointed out that V1 cells are the starting point of visual cortex information processing, generally divided into simple cells and complex cells. Simple cells have a small receptive field and are rectangular in shape, with a central region elongated and a parallel antagonist region on one or both sides of the central region [6]. Simple cells have orientation selectivity, which is almost unresponsive
to diffuse light stimulation, whereas for specific strip stimuli, the response of simple cells varies with the direction of stimulation [7]. The experimental results obtained by many scholars through models show that the cells with orientation selection characteristics are not only responsible for dealing with the orientation information of the edge contours, but also play an important role in the perception and processing of motion signals [8]. This orientation selectivity is basic biological characteristic of V1 cells and plays an important role in edge detection [9]. Complex cells are integrated by simple cells, so the receptive field is larger than the simple cells and the structure is also more complex, as shown in Figure 1.

Based on the dynamic and static characteristics of V1 cells, this paper establishes a V1 cell model with orientation selection for edge detection, and compares with two methods of commonly used edge detection operator Canny and Log. The simulation results show that the cell model has better detection performance and robustness, and the fusion of biological vision and computer vision has been explored.

![Figure 1. The receptive field.](image)

2. The model

2.1. Modeling

Some scholars use Gabor function model, wavelet model or derivative model to simulate the receptive fields of simple cell. [10] Many linear systems and pulse response techniques show that the V1 simple cell is similar to the linear filtering phase of the spatiotemporal energy model [11]. Therefore, this paper uses the spatiotemporal filters of the spatiotemporal energy model to simulate the receptive fields of simple cell, and the spatiotemporal filters are decomposed into separable filters in space and time in order to achieve low computational complexity. In fact, the receptive field of simple cell is the same as the Gabor kernel, and the exponential decay function is a good way to simulate the orientation of cells in time. So spatial component of the filter is described by Gabor filters \( S \) and temporal component by an exponential decay function \( T \). Given the spatial and temporal frequencies \( f_s \) and \( f_t \), of a receptive field, we define the spatial filter and the time filter by:

\[
S (x, y) = e^{-\left[-\frac{(x - x')^2}{2\sigma^2}\right]} e^{j2\pi f_s \cos(\theta) x + f_t \sin(\theta) y}
\]

\[
T (t) = e^{-\left[-\frac{t}{\tau}\right]} e^{j2\pi f_t t}
\]

Where \( \sigma \) and \( \tau \) define the spatial and temporal scales. And \( \theta \) define the preferred orientation of a simple cell. A preferred velocity \( V_c \) related to the frequencies by the relation:

\[
V_c = f_s / f_t
\]

Through the Euler formula, equation (1) and (2) can be rewritten as:

\[
S (x, y) = e^{-\left[-\frac{(x - x')^2}{2\sigma^2}\right]} \left(\cos 2\pi \left(f_s \cos \theta \right) x + f_t \sin \theta \left(\cos \theta \right) y \right) + j \sin 2\pi \left(f_s \cos \theta \right) x + f_t \sin (\theta) \left(\cos \theta \right) y \right)
\]
\[ T(t) = e^{\frac{j\pi}{4}} \left[ \cos(2\pi f_1 t) + j\sin(2\pi f_1 t) \right] \]  \hspace{1cm} (5)

Denoting the real and imaginary components of the filters \(s\) and \(r\) as \(s_r, T_r\) and \(s_i, T_i\) respectively:

\[ S(X,Y) = S_r(X,Y) + jS_i(X,Y) \]  \hspace{1cm} (6)

\[ T(t) = T_r(t) + jT_i(t) \]  \hspace{1cm} (7)

We introduce the odd and even spatiotemporal filters defined as follows:

\[ G_o(X,Y,t) = S_r(X,Y)T_r(t) + S_i(X,Y)T_i(t) \]  \hspace{1cm} (8)

\[ G_e(X,Y,t) = S_r(X,Y)T_r(t) - S_i(X,Y)T_i(t) \]  \hspace{1cm} (9)

The luminance and contrast normalization [10] are implemented as a input \(p(x,y,t)\) pre-processing stage:

\[ P'(x,y,t) = P(x,y,t) - \frac{Lum(t)}{Con(t)} \]  \hspace{1cm} (10)

\(Lum(t)\) and \(Con(t)\) are defined as:

\[ Lum(t) = \sum_i \sum_j P(x,y,t) \]  \hspace{1cm} (11)

\[ Con(t) = \sqrt{\sum_i \sum_j (P(x,y,t) - Luc(t))^2} \]  \hspace{1cm} (12)

We define the response of simple cells with a preferred orientation of contrast sensitivity in the spatial domain, with a preferred velocity:

\[ R_{oe}(x,y,t) = \left[ G_{oe}(X,Y,t) \ast P'(x,y,t) \right] \]  \hspace{1cm} (13)

The complex cells are not a simple superposition of simple cells. it exhibits a strong non-linear property[11]. Many scholars have studied the integration of simple cells. The normalized model [12] can explain the saturation characteristics of cell responses. Energy model [13] can achieve the phase invariance of cell responses. Max model [14] can achieve the scale and local translation invariance. However, the above mechanisms only simulate the partial characteristics of complex cells, without the effective integration of simple cell responses. Based on the existing biological experiment data and models, this paper aims to achieve edge detection by integrating simple cell responses by the following ways:

The half-wave rectification (HR) [10] is used for nonlinear processing:

\[ R_{oe}'(x,y,t) = HR \left[ R_{oe}(x,y,t) \right] \]  \hspace{1cm} (14)

The integration of the above responses is as follows:

\[ E(x,y,t) = \sqrt{R_o(x,y,t)^2 + R_e(x,y,t)^2} \]  \hspace{1cm} (15)

Followed by normalization: the final V1response is defined by:

\[ E_v(x,y,t) = \frac{E(x,y,t)}{\sum_{x,y} E^*(x,y,t) + \varepsilon} \]  \hspace{1cm} (16)
Where $0 \leq \varepsilon \leq 1$ is constant. $w$ are weights, $k_{\text{tuned}}$ and $n$ are the key parameters. When $k_{\text{tuned}} = 0, n = 1$, this equation represents a standard normalization.

According to the experimental simulation results, when $k_{\text{tuned}} = 0, n = 2, w = 1$, the cell model can best fit the biological characteristics.

$$E_{V1}(x, y, t) = \frac{E^2(x, y, t)}{\sum_{\theta} E^2(x, y, t, \theta) + \varepsilon}$$

(17)

2.2. Model analysis

In order to verify the biological characteristics of the V1 cell model, the response of the model to the input stimulus is compared with the biological data. Figure 2(a) and Figure 2(b) show that stimuli are drifting gratings, and plaid patterns composed of two gratings.

![Figure 2. Direction tuning curves.](image)

From Figure 2(c)-2(f), the responses of model are basically consistent with the biological experiment results. The direction tuning curve for gratings is unimodal, but the direction tuning curve for plaid shows two distinct lobes. Each lobe corresponds to one of the plaid’s component gratings. The above experimental results show that the V1 cell model can sense the direction of the stimulus and has selectivity and the universality of biology. So it has laid a good biological foundation for the edge detection experiment.

3. Experiment

In order to verify the effectiveness of the V1 cell model, the edge detection experiment is performed using the MATLAB R2014b software. All images in the experiment are from the USF (University of South Florida) image edge detection database.

3.1. Parameters analysis

It is found that the size and quantity of spatial filters have a great influence on the edge detection of natural images. The experiment result is shown in figure 3(c)-(h) with only changing the size or number of filters (the other parameters constant). Figure 3(a) and 3(b) are the original image and the ground truth.

From the Figure 3, it can be seen that as the size of the spatial filters increases, the edge becomes blurred and even some fine edges overlapping or disappear. When the size of filters is $3 \times 3$, the edge detection performance is best. When the number of filters is 4, some edges can’t be detected. When the number of filters is 6 and 12, the edge detection performance is consistent with the Ground truth. When the number of filters exceeds 6, continuing to increase the number of filters can’t further enhance the edge detection performance, but it will increase the amount of computation. Based on the above analysis, we can see that the edge detection effect of the image is optimal when the size of the spatial filter is $3 \times 3$ and the number is 6.
3.2. Results and analysis

In this paper, the edge detection of all the images in the USF image edge detection database is carried out and compared with the classical edge detection operator Canny and Log operator in order to demonstrate the performance of the model. Three images are selected for demonstration analysis. As shown in Figure 4, the first row is the original image and the second is the ground truth in the USF database, and the third row to the fifth row are the Canny operator, the Log operator and the edge detection result of the model, respectively.

It can be seen from Figure 4, compared with the Ground truth, the results of Canny operator have been more pseudo-edge, so it can’t be used to identify the target. Compared with the results of our model, Log operator has the same detection performance. And both of them can detect edge accurately. But our model contains fewer background pseudo-edges and has a better detection performance. The reason can be concluded that the operator represented by Canny and Log depends...
on the local calculation of the linear filter and uses the bottom feature (the image pixel gray value) to
detect the edge of the image, which amplifies some small and unimportant edges. In this paper, the
two-dimensional Gabor function and the exponential decay function are combined effectively so that
our model has the direction selection characteristics. And the direction (primary feature) is not only
higher than the pixel and gray (underlying features), but also lower than the edge (high-level semantic
features), which is the most useful form to understand the high-level semantic features (edge).
Therefore, the model can detect the edge of the image by using the direction feature, and can achieve
better results.

In order to further demonstrate the anti-noise and robustness of our model, after adding Gaussian
noise (mean 0, variance 0.01), Poisson noise and Spot noise (density 0.05), the results shown in the
Figure 5.

As shown in Figure 5, after adding three different kinds of noise, the original image becomes
slightly blurred. Canny operator is very sensitive to three kinds of noise, so a large number of noise
edges appear in its results. After adding Gaussian noise, Log operator will detect a lot of noise edge
and can’t extract the target edge accurately. Under the Poisson and speckle noise conditions, although
the target shape can be identified, more noise is introduced. For the images with three kinds of noise
respectively, the results of our model are almost no noise. It can detect edge accurately and has
stronger the anti-noise performance and robustness. Noise can significantly affect the gray value of the
image, as a result that the Canny and Log operators can’t effectively distinguish the edge and noise.
Our edge detection model is based on the biometric visual mechanism, which uses the direction of the
characteristics to detect the edge of the image. It can be seen from the equation (15) and (16) that the
model output is an effective integration of a simple cell, which can combine local information to
enhance its ability to distinguish the main body edge and background noise in the image.

For quantitative analysis, this paper calculates the peak signal-to-noise ratio (PSNR) corresponding
to each edge algorithm, as shown in Table 1.

| Algorithm | Gaussian | Poison | Spot   |
|-----------|----------|--------|--------|
| Canny     | 17.16    | 19.98  | 18.86  |
| Log       | 18.85    | 21.39  | 22.10  |
| ours      | 22.48    | 25.07  | 27.25  |
As can be seen from Table 1, compared with Canny operator and Log operator, the proposed model is more effective and has a better robustness.

4. Conclusion
Based on the dynamic and static characteristics of V1 cells, this paper establishes a V1 cell model to detect the natural images, which makes a helpful exploration to the fusion of biological vision and computer vision. Simulation results show that the V1 model can basically fit the biological data and has the universality of biology. The influence of the size and number of the spatial filter on the edge detection of the natural image is analyzed by the control variable method, and the optimal parameters are finally determined. What’s more, compared with the traditional edge detection operators, the proposed model is more effective and has a better robustness. This paper constructs a biological vision model and applies it to image processing. The next stage work will continue to establish a more perfect visual cortex perception model in order to achieve better information representation ability, which will provide more new methods and ideas for the fusion of biological vision and computer vision.

Acknowledgments
This work was supported by NSFC (6137217, 61379104).

References
[1] Namboodiri V M K, Huertas M A and Monk K J 2015. Visually cued action timing in the primary visual cortex. J. Neuron. 86(1) 319-330
[2] Gao S, Peng Y and Guo H 2014. Texture analysis and classification of ultrasound liver images. J. Bio-medical materials and engineering. 24(1) 1209-16
[3] Alberto C, Pietro C and Kiran M K 2004. Calculation of tensile and flexural strength of disordered materials using fractional calculus. J. Chaos, Solitions and Fractals. 37(1) 17-25
[4] Hinton G E, Salakhutdinov R R 2006. Reducing the dimensionality of data with neural networks. J. Science. 313(5786) 504-507.
[5] Hubel D H and Wiesel T N 1959. Receptive fields of single neurons in the cat's striate cortex. J. Journal of Physiology. 148(3) 574-591
[6] Zou H Z, Xu Y L, Ma S P, Li S and Zhang W D 2016. Neuron-filtering feature extraction based on LNP model. J. Journal of Image and Graphics. 21(10) 1376-82
[7] Gao S, Peng Y and Guo H 2014. Texture analysis and classification of ultrasound liver images. J. Bio-medical materials and engineering. 24(1) 1209-16
[8] Burr D C and Thompson P 2011. Motion psychophysics: 1985-2010. J. Vision Research. 51(13) 1431-56
[9] Poo M, Du J and Ip N Y 2016. China Brain Project: Basic Neuroscience, Brain Diseases, and Brain-Inspired Computing. J, Neuron 92(3) 591-96
[10] Xu Y L, Lv C and Ma S P 2017. Motion Feature Extraction of Random-Dot Video Sequences with the Visual Cortex Mechanism, J. Journal of University of Electronic Science and Technology of China. 46(4) 630-635
[11] Yi Z and Bauer P H 2016. Spatiotemporal energy demand models for electric vehicles. J. IEEE Transactions on Vehicular Technology 65(3) 1030-1042.
[12] Kasabov N 2015. Integrative computational neurogenetic modelling. J. Brain Mapping. 667-674
[13] Rust N C, Schwartz O and Movshon J A 2005. Spatiotemporal elements of macaque v1 receptive fields. J. Neuron. 46(6) 945-956
[14] Bradley D and Goyal M S 2008. Velocity computation in the primate visual system. J. Nature Reviews Neuroscience. 9(9) 686-695