

A New Biological Visual Cognitive Behavioural Modeling for Video Energy Computing

Li Zuojin, Chen Liukui

College of Electrical and Information Engineering
Chongqing University of Science and Technology
Chongqing, China, 401331
cqstlzj@sina.cn

Ren Zhiyong

Department of Communication
Chongqing College of Electronic Engineering
Chongqing, China, 401331
cqulzj@gmail.com

Sreenivas Sremath Tirumala

Department of Computing
Faculty of Creative Industries and Business
Unitec of Institute and Technology
Auckland, Newzealand, 1142
zli2@unitec.ac.nz

Abstract—As we all know, human vision is quite sensitive to abnormal behaviors, which is attributed to the discharging of the receptor cells in the brain visual cortex and the ensuing bioelectrical energy features. Inspired by this biological nature, this paper constructs a computing model to describe video dynamic energy, which can further improve the perception of machine vision to abnormal behaviors. Experiments show that this computing model can extract video energy features of abnormal behaviors under complex environment.

Keywords—compute modeling; video energy; visual congenitive;

I. INTRODUCTION

Zone 17 of the brain cortex, namely the primary vision cortex or cortex zone 1, has been understood best among all the brain cortexes. As early as 1962, Hubel etc. proposed for the first time that biological vision cortexes have the structure of function columns [1,2] and found that all cells in vision cortexes share one commonality, the sensibility to directions of external stimuli. However, primary visual cortex only reacts to the input edge within the classic perception scope, which is a small visual area. As the perception objects often exist with a lot of environment information, this traditional perception scope theory is encountered with some limitations when explaining the perception of natural images in larger scopes. Therefore, scientists of biological vision have shifted their focuses onto areas beyond the classic perception scope. And substantial experiments demonstrate that outside of the perception scope of neurons of arena, LGN and visual cortex, there exist areas regulating the reaction of cells, which are much bigger than the traditional perception scope. Within these areas, stimulus of small light spot can not directly arouse the reaction of cells in perception scope, but these spots can suppress the reaction so as to weaken the stimulus of the perception scope itself. To distinguish them with the classic ones, these areas are defined as non-classic scopes, whose effects of inhabitation to discharging are shown in Figure 1. In this figure, (a) is the discharging rate of cells to isolated raster, (b) shows that of cells when raster to the same direction exits outside of the perception scope, (c) occurs when outside the scope there is raster to different directions, and (d) happens when there is no raster in the scope.

The discharging results in Figure 1 illustrate that non-classic perception scope has anisotropic suppression to visual stimuli. When the raster in the environment is to the same direction with the perception objects, the suppression effect is obvious; when the raster is vertical to the direction of the known objects, the suppression is weak. It is the feature of suppression that enables the vision to perceive significant features in a complicated natural scene. As shown in Figure 2, thanks to the suppression of the non-classic scope, we can easily perceive the small raster to different directions. The non-classic scope enlarges the effective scope for receiving feature information of visual cortical neuron by tens of times, which makes it possible to explain the neuromechanism as how the cortical neuron can integrate the image features within a large scope. This feature is essential to deal with complex image information by machine vision.

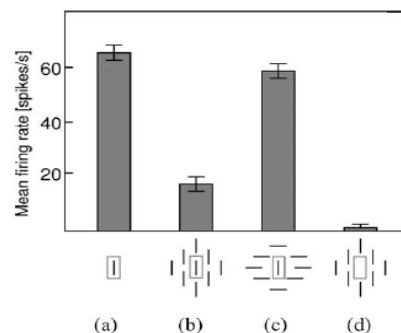


Fig. 1. Inhibition effect of non-classic receptive field on visual cells [3]

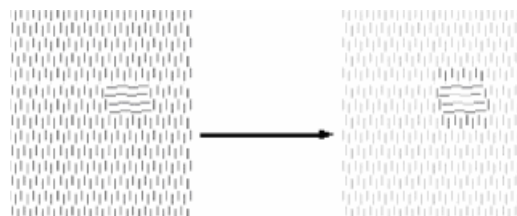


Fig. 2. Lateral inhibition effect of non-classic receptive field

II. COMPUTING MODEL

Based on the sensitivity of simple cells in the vision cortex to raster from different directions, literature [4-6] have proposed to use Gabor function to simulate the directional selectivity of cells, and applied it in selectivity testing of image edges [7-10], whose directivity is mainly represented with the rotation of coordinates in the functions. On this basis, this paper constructs an energy computing model as shown in Figure 3.

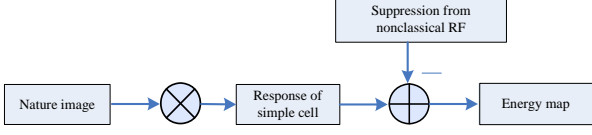


Fig. 3. Image energy computing model inspired by V1 function features

Explanation of the model and its mathematical description:

Response of simple cell: This module simulates the sensitivity of V1 simple cells to edges with Gabor integral transformation proposed in the previous section, which can well test the edge line of images. The algorithm simulation will be shown later in the experiment section.

Memory unit: shows the memory function of cells, which is expressed with the following integral,

$$V(x, y, t) = \int_0^t \dot{V}_i(x, y, \tau) d\tau \quad (1)$$

Suppression module: as described in the first part, stimuli from the environment can suppress the response of cells within the perception scope, which is useful for the extraction of obvious contour. The non-classic scopes make the suppression degrees different in different directions and positions. This model makes use of the feature of anisotropic suppression of non-classic scopes, which can be explained as when stimuli are to the same direction, the suppression is greatest; when they vertical to each other, the suppression is weakest. The relationship between the suppression degree and direction angles can be expressed with the following cosine equation[11],

$$\Delta_{\Theta, \theta}(x, y, x - x_0, y - y_0) = |\cos(\Theta_{\theta}(x, y) - \Theta_{\theta}(x - x_0, y - y_0))| \quad (2)$$

In (2), $\Theta_{\theta}(x, y)$ represents the gradient direction angle of point (x, y) , and $\Delta_{\Theta, \theta}(x, y, x - x_0, y - y_0)$ represents the suppression weight between points (x, y) and $(x - x_0, y - y_0)$ caused by directional difference. Beside the impact of directions, the suppression is also related to the distance between cells, which is shown with the Gaussian error function,

$$w_{\sigma}(x, y) = \frac{H(\text{DoG}_{\sigma}(x, y))}{\|H(\text{DoG}_{\sigma})\|} \quad (3)$$

This, $\text{DoG}_{\sigma}(x, y) = \frac{1}{2\pi(4\sigma)^2} \exp\left(-\frac{x^2 + y^2}{2(4\sigma)^2}\right) - \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$

and $H(z) = \begin{cases} z & z \geq 0 \\ 0 & z < 0 \end{cases}$ represent the first order norm of

phasor, $w_{\sigma}(x, y)$ represents the distance suppression weight of point (x, y) to the origin. The value of the parameter σ will be given later in the experiment section.

Equation (2) and (3) can get the total suppression volume of environment to the given stimulus point,

$$T_o = \sum_i M_i \Delta_{\Theta, \theta} w_{\sigma} \quad (4)$$

Here, i represents the environment of the interest point o , M_i represents the excitement response of environment stimuli, or the gray value of pixels in images.

Energy map module: based on the amplitude and range, the video energy can be calculated as follows,

$$E_o(n) = \sum_{i=1}^W \sum_{j=1}^H w_{i,j}(n) \bullet v_{i,j}^2(n) \quad (5)$$

In (5), the parameter $E_o(n)$ represents energy value, $v_{i,j}(n)$ represents the dynamic features in the position (i, j) of the image n (with the width of W , and height of H), and $w_{i,j}(n)$ is the weight of the dynamic features, represented with the gradient direction of the Module 2.

III. EXPERIMENT

In order to prove the effectiveness of this model, experiments use two scenes as the objects for data processing, as shown in Figure 4, in which (a) is a single person scene, (b) is an abnormal scene. Evidently with biological facts, visual cells do not receive intense stimulus when seeing Figure 4(a), while seeing (b), human visual cells will show great stimulatory effect because of its abnormality, releasing much more bioelectrical energy and being more sensitive to the latter. On this basis, this paper proposes a computer energy representation model with characteristics of human visual effect. Figure 5(a) and (b) shows the energy amplitudes of cells for Figure 4 (a) and (b) respectively. As we can see, scene (b) shows greater computing energy value, consistent to the intensity of stimulus to visual cells. Similarly, when we analyze its directional distribution, we get the results shown in Figure 6. It shows, when we notice abnormal behaviors, the directions of the scene become scattered, while in a quiet scene, the directions are concentrated. This surprisingly coincide with the directional distribution of human vision; that is, in scene (a), human vision will follow the movement of the single person, while in scene (b), human vision will be attracted by different

stimuli and show scattered distribution. Our computing model get the results that are consistent to the response of human vision.



Fig. 4. The real scene for experiment. (a) Single person scene, (b) abnormal scene.

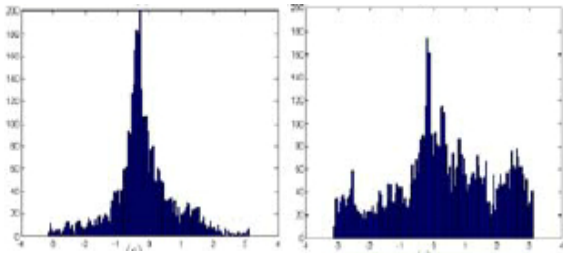


Fig. 5. The energy amplitudes for Fig 4 (a) and (b) respectively

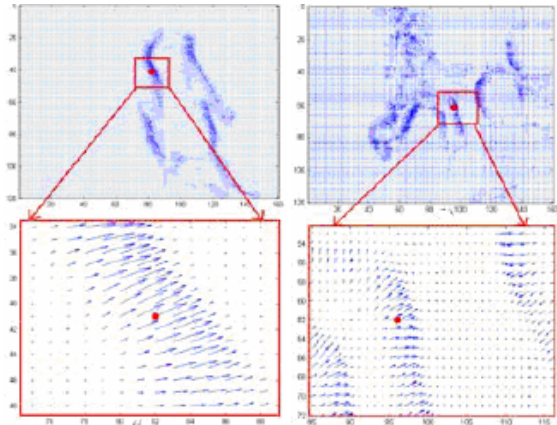


Fig. 6. The energy directional distribution for fig 4 (a) and (b) respectively.

IV. CONCLUSION

Human visual system has strong perceptivity, so the simulation of biological visual features for strengthening computer vision processing ability has attracted high attention of scholars from home and abroad. This paper, inspired by the strengthened stimulus of abnormal behaviors to visual cells,

constructs a video energy computing model based on visual features. Experiments prove that this model and algorithm possess biological visual features in representing energy in complex scenes, which can add great value to the methodology of machine vision and its application in engineering.

ACKNOWLEDGMENT

This work was supported in part by Scientific and Technological Research Program of Chongqing (Grant No. KJ131423, No. KJ131422 and KJ132206), Doctor Special Project from Chongqing University of Science and Technology (Grant No. CK2011B05 and CK2011B09) and Natural Science Foundation of Chongqing (Grant No. cstcjjA40041).

REFERENCES

- [1] D. Hubel and T. Wiesel, "Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex," *Journal of Physiology*, London, Vol.160:106-154, 1962
- [2] S. Zeki, *A Vision of the Brain*. Oxford, Blackwell Scientific Publications, 1993.
- [3] H. C. Nothdurft, J. Gallant, Response modulation by texture surround in primate area V1: Correlate of "popout" under anesthesia, *Visual Neuroscience*, 1999, 16(1), 15-34.
- [4] C. Chang and S. Chatterjee, "Ranging through gabor logons-a consistent, hierarchical approach," *IEEE Transactions on Neural Networks*, 1993, vol. 4, no. 5:827-843.
- [5] T. C. Folsom and R. B. Pinter, "Primitive features by steering, quadrature, and scale," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, vol. 20, no. 11:1161-1173.
- [6] M. G. Milanova, A.S. Elmaghraby, M. P. Wachowiak, and A. Campilho, "A computational model of visual cortex receptive fields using spatio-temporal filters," in *proceedings of the 2000 IEEE/EMBS Conference on Information Technology Applications in biomedicine*, 2000:129-134.
- [7] Q. L. Tang, N. Sang and T. X. Zhang, "Extraction of salient contours from cluttered scenes," *Pattern Recognition*, 2007, vol. 40:3100-3109.
- [8] W. T. Hang, L. C. Jiao and J. H. Jia, "Modeling contextual modulation in the primary visual cortex," *Neural Networks*, 2008, Vol. 21:1182-1196.
- [9] C. Grigorescu, N. Petkov and M. A. Westenberg, "Contour and boundary detection improved by surround suppression of texture edges," *Image and vision computing*, 2004, vol. 22, no. 8:609-622.
- [10] Fang Fang, Huseyin Boyaci, and Daniel Kersten. Border Ownership Selectivity in Human Early Visual Cortex and its Modulation by Attention, *The Journal of Neuroscience*, 2009, 29(2): 460-465
- [11] C. Grigorescu, N. Petkov and M. A. Westenberg, "Contour and boundary detection improved by surround suppression of texture edges," *Image and vision computing*, 2004, vol. 22, no. 8, pp.609-622.