A Scale-Changeable Image Analysis Method

Hui Wei, Bo Lang, Qing-song Zuo

Laboratory of Cognitive Model and Algorithm, School of Computer Science and
Technology, Fudan University, Shanghai, China

weihui@fudan.edu.cn, blang2004@gmail.com, qingsong.zuo@gmail.com

Abstract. The biological vision system is far more efficient than machine
vision system. This is due to the former has rich neural layers for representation
and process. In order to obtain a non-task-dependent image representation
schema, the early phase of neural vision mechanism is worth simulating. We
design a neural model to simulate non-classical receptive field of ganglion cell
and its local feedback control circuit, and find it can represent image, beyond
pixel level, self-adaptively and regularly. The experimental results prove this
method can represent image faithfully with low cost, and can produce a com-
pact and abstract approximation to facilitate successive image segmentation as
well as integration operation. This representation schema is good at extracting
spatial relationship from different components of image, thus it can be applied
to formalize image semantics. Further it can be applied to object recognition or
image classification tasks in future.

Keywords: Neurons, non-classical receptive field, image representation,
feedback control circuit

1   Introduction

The function of any representation scheme is to capture the essential features of
an image and make those features accessible at higher processing layers. The
target of image representation is to discover what objects are in an image, and how
about their spatial relations. Human vision system generates an internal representation
for the external stimulus and requires an internal data structure recording the spatial
distribution of stimulus. Such a fundamental data structure is also required by
computer vision for recording the image. A better representation should characterize
an image by the spatial relationships between them such that a later processing can be
adapted to preserver or enhance these relationships.

In order to achieve this goal, a bio-inspired image representation model based on
non-classical receptive field (nCRF) is constructed, in which every nCRF denotes a
unit of visual information processing, basic, structural and functional. Ganglion cell
(GC) merges all stimuli occurring in its RF and reports a composite message upwards
for further processing. By means of these dense and regular RFs, many GCs can
realize a general representation for any outside stimuli.

Image representation is of primary importance for developing effective image
processing and analysis techniques. However, due to lack of a complete
understanding of the human vision system and the existence of an infinite number of
probable image patterns, it is extremely difficult to construct a universal model for
various natural scene images. Facing this dilemma, researchers in this field have

resorted to designing separate models for different applications. Some mechanisms proposed for image representation range from color histogram to feature statistics, from spatial frequency based to region based, and from color detection to topology detection. For a more extensive review of image representation techniques, see [1], [2], [3], [4]. Another image representation scheme using a set of block pattern models was constructed in which each model a small image block is satisfying certain intensity variation constraints. Although many representation schemes have been proposed, local image patterns was characterized by their statistical properties (i.e. pixels) in most of them. Those operations only differentiating and labeling pixels, this is far away from grasping semantic of an image. As we known, an image is an array of pixels, and this kind of discrete and separate structure cannot manifest semantic organization of image in-depth. Assembled or clustered pixels are more meaningful and effective than single pixel. So a novel representation fashion that can make pixels clustered into some sets or regions is required. Our formulation is unique in that each RF represents explicitly a basic image unit and is described by block-grained. Each block is a compact unit including semantic and every of them should be perfectly divisible units. There is a vivid metaphor describing it, holding a dozen of small tones must be much easier than holding 10-thousands grains of sands with same weight. Integrating pixels of image and use a set to pocket them is an indispensible step in image understanding. It provided a heuristic approach that integrating properly and reducing the load of memory will greatly facilitate semantic emerging.

The arrangement of this paper is as follows. In section II, we formalize the design of the image representation model based on nCRF. Some experiment results are presented in Section III. Finally, in Section IV, we conclude the paper by summarizing the main result and suggesting possible future investigations.

2 Formulations of The Multi-layer Computational Model Based on nCRF

2.1 A Computational Model for RF Self-adaptive Adjusting Dynamically

Neurophysiological shows that the horizontal cell (HC) with a wide range of electrical synaptic connections, it can receive inputs from large areas of receptor cells (RCs). When a disinhibitory nCRF are solely and fully stimulated by grating patches with low spatial frequency, large area horizontal cells can be simultaneously activated. It affects the activities of RCs by means of spatial summation and feedback and then elicits the (GC) responses to the stimuli within the disinhibitory nCRF through bipolar cells (BCs). Moreover, the amacrine cell (AC) connects many of the nearby GCs in the horizontal direction through its extensive dendritic branches. ACs also inter-connect with each other. The breadth of AC's connection exceeds far beyond GC's CRF surround, So AC is properly related to the formation of nCRF with a wide range. Thus, HC and AC play the role of information integration in the outer and inner plexiform layer respectively. GC receive inputs from many neurons in the outer and

inner plexiform layer, hence HC and AC are properly connected with the formation of nCRF of retinal GC. Figure 1 shows the details of computational model. According to the computational model show in Fig. 1. The RF's size can self-adaptive adjusting according to the color changing of region is uniform or sharp.
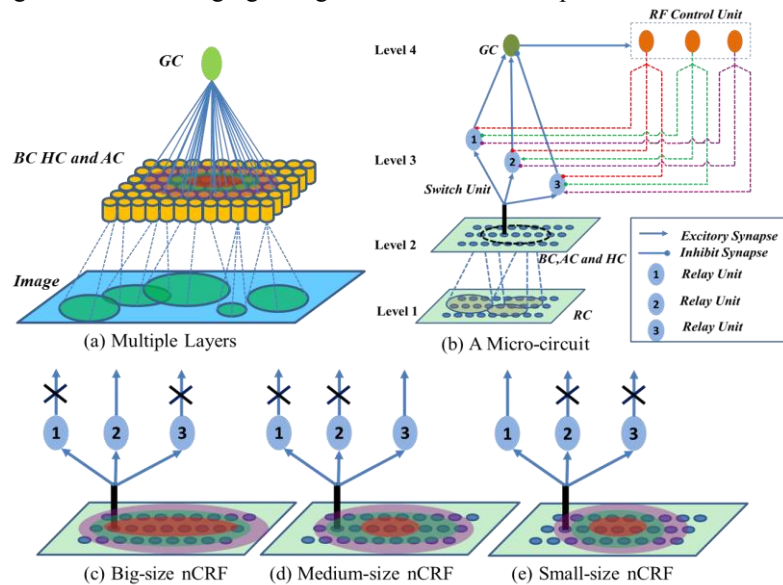

Fig. 1. A multi-layer neural architecture and how a GC adjusts its nCRF dynamically.

(a)is a multiple layers architecture. For the sake of clarity, only one GC and its nCRF is drawn. A nCRF has three rings with positive, negative and positive weight values. (b) This is a neural circuit of nCRF adjusting dynamically, which is a small function unit of (a) in detail. Because a neuron can change its output projection direction according to changing stimulus, nCRF can dynamically change its size. This can be realized by three relay neurons and three switch neurons. A switch neuron imposes its backward control on three relay neurons, and selectively permits only one relay outputting its signal upwards to GC. The relay neuron may have a chance to join one of three different rings of nCRF(c)-(d). With the different switch is turning on ,the same neuron may exclusively participate in forming one of rings of a nCRF. Then a size-changing nCRF come into being. (c) is big one, (d) is middle one and (e) is small one.

3. Experiments Results

A series of experiments is performed, from several single attributes to the whole function one-by-one, to test the efficiency of nCRF based representation model.

3.1. The Relationship between Entropy and Number of RF

Information entropy is a key concept in information theory. If this concept is applied in image description, it can measure the extent of color change in an image. If we define the quantity of information contained in an image as the number of blocks it contains, then according to the equation of entropy, we get:

$$I[M] = \left(\sum_{i=1}^{n} -p(m_i) \log_2(p(m_i)) \right) = E\left[-\log_2 p(m_i) \right]$$. where M is a set of

classes of blocks with different sizes, M = {m1, m2, ……, mn}, and each class has multiple occurrences, meaning there can be a number of blocks of the same size. A set {p(mi)} defines the numerical likelihood that all classes occur. I[M] represents the information entropy of an image. In terms of the same strategies used in two previous experiments, we calculated information entropies of 300 images, and sorted them into a sequence based on these values.
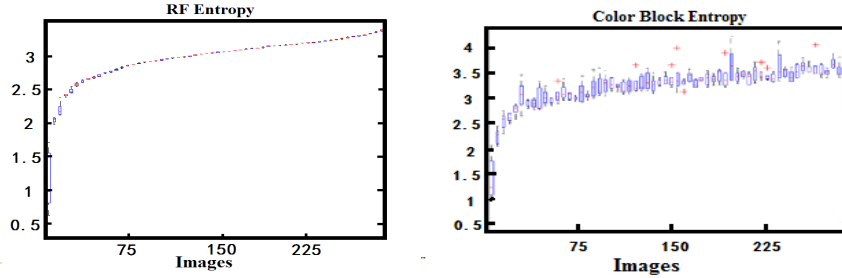

Fig. 2. RF-based entropy and block-based entropy

3.2 Multiscale Image Representation Using nCRF

There exists hierarchical structure among biological tissue of vision, characteristics of external object and the image process. Therefore, the image feature must be described clearly by vision image with multi-scale. Utilizing wavelet transform, a unified mathematic description of the scale of vision image can be achieved with multiscale or multiresolution analysis. The wavelet $\psi(x)$ is stretched as $\psi(x/a)$ (*a* denote scale) to construct a set of primary functions. When the value of *a* arises, the primary function can represent the global feature of image, on the contrary, while it descends, it can be used to search for the detailed information of image.

From the pattern of nCRF, the equivalent functional form of single GC or LGN cell for performing calculation in local scope, and the exhaustive distribution of GC and LGN array on the whole vision, it can be concluded that the process of GC and LGN array for image is very similar to the wavelet transform. In fact, DOG function, which is widely used by researcher, is a wavelet function itself, too. A distinctive feature of wavelet transform is the multiscale analysis, which can use different observing scale to detect and represent different size characteristics. The RF of GC and LGN are capable of dynamically changed, it realized the effective extraction and representation for local characteristics of stimulation. For example, the large continuous areas are suitable for being represented by large-sized RF, while the details of image are

suitable for being represented by small-sized RF. In this regard, the nervous system shows the adaptability and suitability of the structure and functionality. To explore the mathematical essence of the physiological mechanism, wavelet transform is a very appropriate tool for description. Fig. 3 illustrates the close association between image processing result from the nCRF mechanism and wavelet transformation

The experiment result shows that:

(1) With the increasing of level of wavelet transformation, the details of image are gradually faded and the regional characteristic of image emerges simultaneously. The characteristic in these space positions can be exactly rep-resented by the lager-size RF.

(2) The details of image can be exactly represented by the small-size RF.

(3) Multi-scale analysis come from wavelet transform can be represented by Multi-size RF.

It can be found that nCRF is an image analysis using neural computation actually, of which function is similar to wavelet transform. This method is credible because its foundation of mathematic is derived from wavelet transform. Although computation ability may not exceed wavelet transform, the representation based on neuron can achieve the synchronous enhance effect in a certain scope by means of synchronous oscillations, and make them prominent from the whole group of neurons. All of these may be the mechanism of neurons, of which can realize the image integration and segmentation.
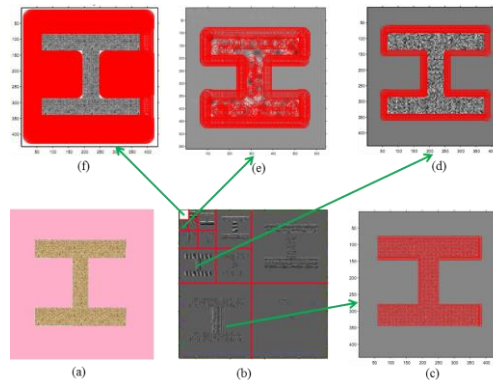

Fig. 3. Multiscale image representation using nCRF. The red circle denotes the RF with different size. (a) is original image. (b) is an image after wavelet transform.(c) is the distribution of minimum-size RF, which able to represent the details of face in the image. (d) is the distribution of medium-size RF. (e) is the distribution of large-size RF. (f) is the distribution of full-size RF, which able to represent the region with homogeneous color in the image.

4. Discussion

The feasibility for image representation based on nCRF is studied in this work. The real effect of nCRF is newly thought as increasing the encoding efficiency and lowering the redundancy of cells. Visual system can be regarded as a system of information processing and encoding according to information theory [12]. If visual

system takes adaptation to natural stimulus as its evolutional goal, it can certainly employ the most effective encoding method for natural images. Since natural images have high relativity, the visual information inputting to RCs is largely redundant [13]. They found that the modulation of nCRF increases the selectivity of the responses of V1 neurons in and the sparsity of the response distribution of groups of neurons. The narrow effective bandwidth of the response curve of single neuron does not decrease the amount of information. Therefore, they thought that the modulation of nCRF help improve the encoding efficiency of visual information. For this purpose, two measures should be taken: first, increase the encoding efficiency of single cell and fully utilized the dynamic characteristics of the cell itself; second, reduce as much as possibly the redundancy in cells possibly and use as few cells as possible to transfer information, which remains the principal problem for our future work.

References

1. Eauqueue J, Boujemaa N.: Region-based Image Retrieval: Fast Coarse Segmentation and Fine Color Description. J. Vision Languages and Computing. 15(1):69-95 (2004)
2. Deng Y, Manjunath B S, Kenney C.: An Efficient Color Representation For Image Retrieval. , J. IEEE Transaction on Image Processing. 10(1):140-147 (2001)
3. Ediz Saykol, Ugur Gudukbay, Ozgur Ulusoy.: A Histogram Based Approach For Object Based Query By Shape And Color In Image And Video Database. J. Image and Vision Computing, 23, 1170-1180, 2005. (1999)
4. Jeong S, Won C S, Gray R M.: Image Retrieval Using Color Histograms Generated By Gauss Mixture Vector Quantization. J. Computer Vision and Image Understanding, 9(1-3):44-46 (2004)
5. Yao Wang and Sanjit K. Mitra.: Image Representation Using Block Pattern Models And Its Image Processing Applications. J. IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol 15, NO. 4 (1993)
6. Allman J, Miezin F, McGuiness E.: Stimulus Specific Responses From Beyond The Classical Receptive Field: Neurophysiological Mechanisms For Local Global Comparisons On Visual Neurons. J. Annu Rev Neurosci, 8:407-430 (1985)
7. Fitzoatrick D.: Seeing Beyond The Receptive Field In Primary Visual Cortex. J. Current Opinion in Neurobiology, 10:438-443 (2000)
8. Sillito AM, Grieve KL, Jones HE, Cudeiro J, Davis J.: Visual Cortical Mechanisms Detecting Local Fical Orientation Discontinuities. J. Nature 378:492-496 (1995)
9. Wilson HR, Richards WA.: Curvature and Separation Discrimination at Texture Boundaries. J. Opt Soc Am A 9:1653-1662 (1992)
10. Krieger G, Zetzsche C.:Nonlinear Image Operators For The Evaluation Of Local Intrinsic Dimensionality. J. IEEE Trans Image Process 5:1026-1041 (1996)
11. http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/
12. Zetzsche, C., Röhrbein, F.: Nonlinear And Extra-Classical Receptive Field Properties And The Statistics Of Natural Scenes, J. Network, 12, 331-350 (2001)
13. Ruderman D.L. et al.: Statistics Of Natural Images: Scaling In The Woods, J. Phys. Rev. Lett., 73, 814-817 (1994)