Neurocomputing 164 (2015) 96–111

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A biologically inspired neurocomputing circuit for image representation

Hui Wei, Qingsong Zuo

Laboratory of Cognitive Algorithm Model, Department of Computer Science, Shanghai Key Laboratory of Data Science, Fudan University, Shanghai 200433, China

ARTICLE INFO

Article history:
Received 3 January 2014
Received in revised form
21 January 2015
Accepted 24 January 2015
Available online 24 March 2015

Keywords:
Neural model
Retina
Image representation
Non-classical receptive field
Feedback control circuits
Dynamic self-adjustment mechanism

ABSTRACT

Biological vision systems have evolved over millions of years, resulting in complex neural structures for representation and processing of stimuli. Moreover, biological visual systems are typically far more efficient than current human-made machine vision systems. The present report describes a non-task-dependent image representation scheme that simulates a biological neural vision mechanism in the early visual system. We designed a neural model involving multiple types of computational units to simulate ganglion cells and their non-classical receptive fields, local feedback control circuits and receptive field dynamic self-adjustment mechanisms in the retina. Beyond the pixel level, our model is able to represent images self-adaptively and rapidly. In addition, the improved representation was found to substantially facilitate image segmentation, figure-ground separation, saliency detection, and object recognition. We propose that these improvements arise because the retinal ganglion cells can resize their receptive fields, enabling multi-scale analysis functionality, a neighborhood referring function, and a localized synthetic function. The ganglion cell layer is the starting point for a diverse variety of subsequent visual processing. The extracted features and image presentation by the ganglion cell will be transmitted into high levels of visual system, for many visual tasks such as image segmentation, contour detection, object recognition and so on, the visual representation in the early stage of visual system is universal and independent on visual tasks.

1. Introduction

The present work describes a new type of scheme for image representation is described here. This new schema is based on the simulation of biological processes, because biological visual systems are highly optimized as a result of evolutionary pressures. This system involves modeling retinal ganglion cells (GCs) and their dynamic properties.

Biological visual systems, especially the vision of higher mammals, can adapt to changing environments. A human observer can easily recognition his friend in a crowd. However, the same task becomes extremely challenging for a computer vision system due to the absence of an appropriate representation of the object and context. Some questions are worth to researching: what features are good for a computational method? What features are used by humans for recognition? How the visual system represent the input visual information. In the human visual system, visual information obtained from the eye is first processed by the circuitry of the retina. The output of the retina is conveyed to the lateral geniculate nucleus by the axons of GCs, which form the optic nerve. By analyzing the output of the retina, researchers can obtain the methods of processing and transforming in the retinal receptive fields (RFs). A typical RF of a GC is a circular area on the retina, consisting of a central region and an annular surrounding region. A number of computational studies in the past 20 years have focused on the modeling of GCs: typically with respect to edge or contour detection [1], image enhancement [2], and multi-scale analysis [3]. Some recent studies have applied GC models to represent images [4–6], and have begun to model much more complex functions of a GC RF. Consideration of these studies from the perspective of neural computation reveals that the following two aspects can be improved upon: (a) referring to a greater variety of neurobiological findings to construct a more complete neural-processing cell circuit, and (b) adopting a dynamic receptive field strategy that requires cell-circuits to adjust their RFs self-adaptively. As such, the main purpose of this paper is to design a GC model that imitates neural mechanisms to represent images.

The remainder of this paper is arranged into the following sections. In Section 2, we formalize the design of the image representation model based on the biological mechanisms of the GC and its RF, and present an algorithm of adaptively adjusting the

E-mail addresses: weihui@fudan.edu.cn (H. Wei), 12110240028@fudan.edu.cn (Q. Zuo).

http://dx.doi.org/10.1016/j.neucom.2015.01.078
0925-2312/

size of the RF. Sections 3 and 4 present the models on segmentation and figure-ground separation methods, respectively, applied to input images. Section 5 shows that these representations can effectively promote saliency detection. In Section 6, an experiment for object detection is proposed, and the results show that our image representation scheme is successfully applied to improve object detection. Finally, Section 7 concludes the paper and discusses the future work on computer vision, with emphasis on models inspired by the human visual system.

2. A model based on non-classical receptive field

In the human vision system, GCs are the final stage of retinal information processing, thus GCs and their RFs underlie almost all final information processing functions in the retina. Here, we present a GC-inspired multiple layer model for image representation. The focus is on a non-classical receptive field (nCRF) mechanism.

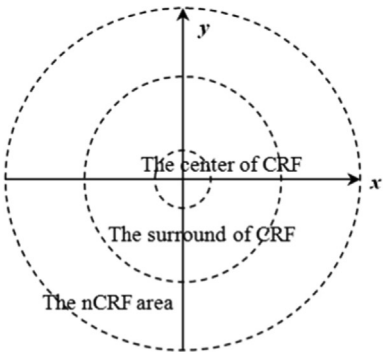
2.1. The anatomical structure of the nCRF

Since the 1960s, numerous studies have established the existence of a large region outside the classical receptive field (CRF) [7–9]. In the outer region, light spot stimuli cannot directly cause a reaction of the cells, but stimuli in this region can modulate the reaction caused by the CRF. This modulation can be facilitatory, inhibitory or disinhibitory [10,11], and this expanded receptive field is denoted as an nCRF. The RF is divided into three parts: the center of CRF, the surround of CRF and the nCRF area, as shown in Fig. 1. Activities in the region can inhibit the antagonistic effect and compensate in part for the loss of low spatial frequency caused by the CRF center-periphery antagonism to some extent. The nCRF plays an important role in representing contour [12], shape [13], and curvature [14,15]. Moreover, it plays an important role in separating figures out of background. The nCRF compensates for the loss of low spatial frequency by adding the output from the surround of CRF. Through its nCRF, a GC expands its information-receiving scope of its CRF, and, undoubtedly, this neural basis makes the GC the capable of integrating image features over a larger scale [16].

Neurophysiology establishes that the ability of horizontal cells (HCs) to integrate information from widely separated receptor cells (RCs) depends on the permeability of large gap junctions. When a disinhibitory nCRF is solely and fully stimulated by grating patches in low spatial frequency, large-scale HCs can be activated simultaneously. This affects the activities of RCs through spatial summation and feedback, and elicits the responses through bipolar cells (BCs). Here, spatial summation is defined as a way of achieving an action potential in a neuron with input from multiple presynaptic cells. Spatial summation is the algebraic summation of potentials from different areas of input, usually on the dendrites. Moreover, amacrine cells (ACs) connect many of the nearby GCs in the horizontal direction through extensive dendritic branches. ACs also interconnect with each other. The span of AC's connection extends far beyond a GC's CRF surround. Therefore, AC is properly related to the formation of an nCRF over a wide range. Thus, HC and AC play the role of information integration in the outer and inner plexiform layer respectively. GCs receive inputs from many neurons in the outer and inner plexiform layer, hence, HCs and ACs are properly connected with the formation of the nCRF of a retinal GC [17].

2.2. Model of a GC adjusting its RF self-adaptively

GCs adjust their RFs according to the visual input. RFs will expand at low contrast and shrink at high contrast. As a result, RFs with a variety of sizes are more efficient for representing an image than RFs of an equivalent size. Therefore, finding a way to apply the nCRF to image understanding (rather than just edge detection, image smoothing or contrast invariance), remains problematic, especially for information integration. An important finding is that a GC RF can be resized to a certain extent according to the properties of a given stimulus [8,18–20]. Smirnakis et al., recorded the spike trains of GCs in the isolated retina of a tiger salamander or rabbit, and found that the visual processing of the GC in the retina is adaptive and adjusts not only to the average illumination but also to both the range of intensity fluctuations and their spatial scale [21]. The study demonstrated a remarkable plasticity in the retinal processing that may contribute to the contrast adaptation of the human vision. Solomon et al. quantified the effect of nCRF stimulation on visually responsive cells in the lateral geniculate nucleus (LGN) of a marmoset through electrophysiological experiments [22]. The effect of the nCRF is referred as extra-classical inhibition (ECI). The study showed that the ECI also contributed to contrast-dependent changes in spatial summation. The researchers found that LGN cells also show contrast-dependent changes in spatial summation, where, on average, the size (radius) of the excitatory CRF at a low contrast is 1.31 times that at a high contrast [22]. Therefore, for the LGN cells, the excitatory summation region was shown to be larger at a lower contrast. For the parvocellular cell, the size of the excitatory region at a low contrast is 1.45 times that at a high contrast. For the magnocellular cell, the ratio of the sizes of the excitatory at low contrast to that at high contrast is 2.37, and, for the koniocellular cell, the ratio is 1.16. The radius of the inhibitory region exhibited less consistent changes. In 2004, Nolt et al. found that spatial summation within their RFs was dependent on the contrast of the stimuli presented, based on extracellular recordings from 69 LGN cells in anesthetized cats [23]. The researchers proposed that this contrast dependency in the retinal GCs results directly from a reduction in the size of the central mechanism due to an increase in contrast. They also pointed out that these properties first arise in the retina and are



$$GC's\ output = \int_{Center} e_1(x,y)I(x,y)dxdy - \int_{Surround} e_2(x,y)I(x,y)dxdy + \int_{nCRF} e_3(x,y)I(x,y)dxdy.$$

where e_1, e_2, e_3 are positive weight functions, and $I(x,y)$ is the intensity of light at (x,y).

Fig. 1. Spatial structure of the CRF and nCRF and their computational relationships.