

Early Visual Processing for Pattern Recognition in Natural Environments

Rosana Matuk Herrera

Department of Computer Science, Facultad de Ciencias Exactas y Naturales,
Universidad de Buenos Aires, Argentina
rmatuk@dc.uba.ar

Abstract. In this article, the performance of hierarchical architectures for computer vision of biological inspiration is analyzed. In particular, the role of the lateral geniculate nucleus (LGN) for the robustness in the recognition of ring shaped patterns under gradual changes in illumination was analyzed. Changes in illumination are common in everyday environments and thus, it is a crucial factor to be considered for the recognition of objects in everyday environments.

1 Introduction

In this article, we analyze a computational model of the visual cortex, and in particular the role of the early visual processing system, for the recognition of patterns under different levels of illumination. This is also a fundamental problem for robotics in uncontrolled environments, because illumination strongly affects the visual perception of the environment by the robots.

2 Biological Background

During visual perception, light entering the eye is detected by the retina, an array of photoreceptors and related cells on the inside of the rear surface of the eye. The cells in the retina encode the light levels at a given location as patterns of electrical activity in neurons called ganglion cells. Output from the ganglion cells travels through neural connections to the lateral geniculate nucleus of the thalamus, or LGN, at the base of each side of the brain. From the LGN, the signals continue to the primary visual cortex, or V1 at the rear of the brain (Fig. 1). The output from V1 goes on to many different higher cortical areas, including areas that underlie object and face processing.

Retinal ganglion cells perform a type of edge detection on the input, responding most strongly to borders between bright and dark areas. The two main types of such neurons are the ON-center and OFF-center. An ON center retinal ganglion cell responds most strongly to a spot of light surrounded by dark, located in a region of the retina called its receptive field, or RF. An OFF-center ganglion cell instead prefers a dark area surrounded by light. The size of the preferred

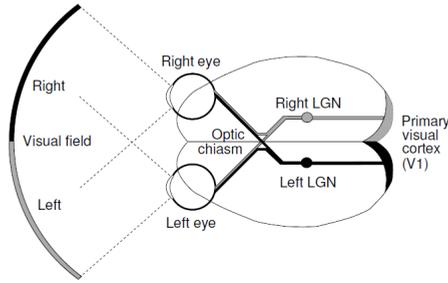


Fig. 1. Human visual pathway (extracted from [1])

spot determines the spatial frequency preference of the neuron; neurons preferring large spots have a low preferred spatial frequency, and vice versa.

Neurons in the LGN have properties similar to retinal ganglion cells, and are also arranged retinotopically, so that nearby LGN cells respond to nearby portions of the retina. The ON-center cells in the retina connect to the ON cells in the LGN, and the OFF cells in the retina connect to the OFF cells in the LGN. Because of this independence, the ON and OFF cells are often described as separate processing channels: the ON channel and the OFF channel.

Like LGN neurons, nearby neurons in V1 also respond to nearby portions of the retina and are selective for spatial frequency. Unlike LGN neurons, most V1 neurons are binocular, responding to some degree to stimuli from either eye, although they usually prefer one eye or the other. They are also selective for the orientation of the stimulus and its direction of movement. In addition, some V1 cells prefer particular color combinations (such as red/green or blue/yellow borders), and disparity (relative positions on the two retinas). V1 neurons respond most strongly to stimuli that match their feature preferences, although they respond to approximate matches as well (see [2] for a review).

3 Computational Maps of the Visual Cortex

The LISSOM model (laterally interconnected synergetically self organizing map; [3] [4] [5] [1]) was designed as a computational map model of V1.

The V1 network in LISSOM is a sheet of $N \times N$ interconnected computational units, or neurons (Fig. 2). Because the focus is on the two-dimensional organization of the cortex, each neuron in V1 corresponds to a vertical column of cells through the six layers of the biological cortex. This columnar organization helps make the problem of simulating such a large number of neurons tractable, and is viable because the cells in a column generally fire in response to the same inputs.

Each cortical neuron receives external input from two types of neurons in the LGN: ON-center and OFF-center. The LGN neurons in turn receive input from a small area of the retina, represented as an $R \times R$ array of photoreceptor cells. The afferent input connections from the retina to LGN and LGN to V1

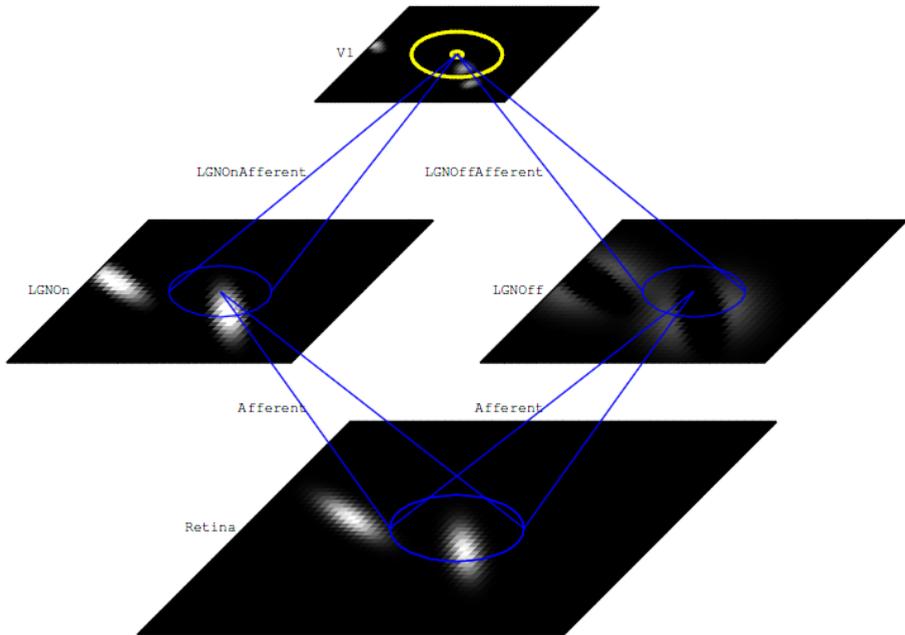


Fig. 2. Architecture of the LISSOM model. The large circle in V1 corresponds to the range of lateral inhibitory connections, and the small circle to the range of lateral excitatory connections, regarding to the neuron situated at the center of the V1 sheet.

are all excitatory. In addition to the afferent connections, each cortical neuron has reciprocal excitatory and inhibitory lateral connections with other neurons. Lateral excitatory connections have a short range, connecting only close neighbors in the map. Lateral inhibitory connections run for long distances, but may be patchy, connecting only selected neurons.

The ON and OFF neurons in the LGN represent the entire pathway from photoreceptor output to the V1 input, including the ON/OFF processing in the retinal ganglion cells and the LGN. Although the ON and OFF neurons are not always physically separated in the biological pathways, for conceptual clarity they are divided into separate channels in LISSOM. Each of these channels is further organized into an $L \times L$ array corresponding to the retinotopic organization of the LGN.

Each neuron develops an initial response as a weighted sum (scalar product) of the activation in its afferent input connections. The lateral interactions between cortical neurons then focus the initial activation pattern into a localized response on the map. After the pattern has stabilized, the connection weights of cortical neurons are modified. As the self-organization progresses, these neurons grow more nonlinear and weak connections die off. The result is a self-organized structure in a dynamic equilibrium with the input.

4 Method

For the experiments, we use Topographica [6] which is a software package for computational modeling of neural maps based on LISSOM, developed by the Institute for Adaptive and Neural Computation at the University of Edinburgh and the Neural Networks Research Group at the University of Texas at Austin.

In order to analyze the role of the ON/OFF channels, two different architectures were tested. The first architecture is the basic LISSOM model (Fig. 2), which has a retina network at its bottom level, ON and OFF networks at its middle level, and a V1 network at its top level. The second architecture, is similar to the the first architecture, but without the ON and OFF networks (Fig. 3).

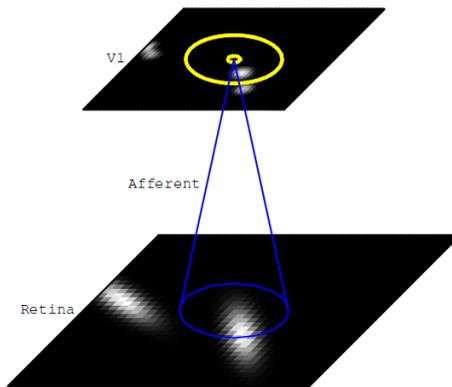


Fig. 3. Architecture without the ON/OFF channels

The following dimensions were used: a 54x54 retina, a 36x36 LGN (composed of one 36x36 OFF channel sheet, and one 36x36 ON channel sheet), and a 48x48 V1. Both architectures were trained presenting 10000 pairs of oriented Gaussian patterns at random angles and positions.

5 Results

The training of the networks lasted around 5 minutes on a computer with an Intel Core i5 processor. The response of the network after the training, when a test pattern was presented, was very fast (less than 1 second).

The orientation Preference/Selectivity maps and Fourier spectrums obtained are shown in Fig. 4. Figure 5 shows the responses of the V1 activity under testing. The test patterns consisted of a ring shaped pattern, increasing its level of smoothing.

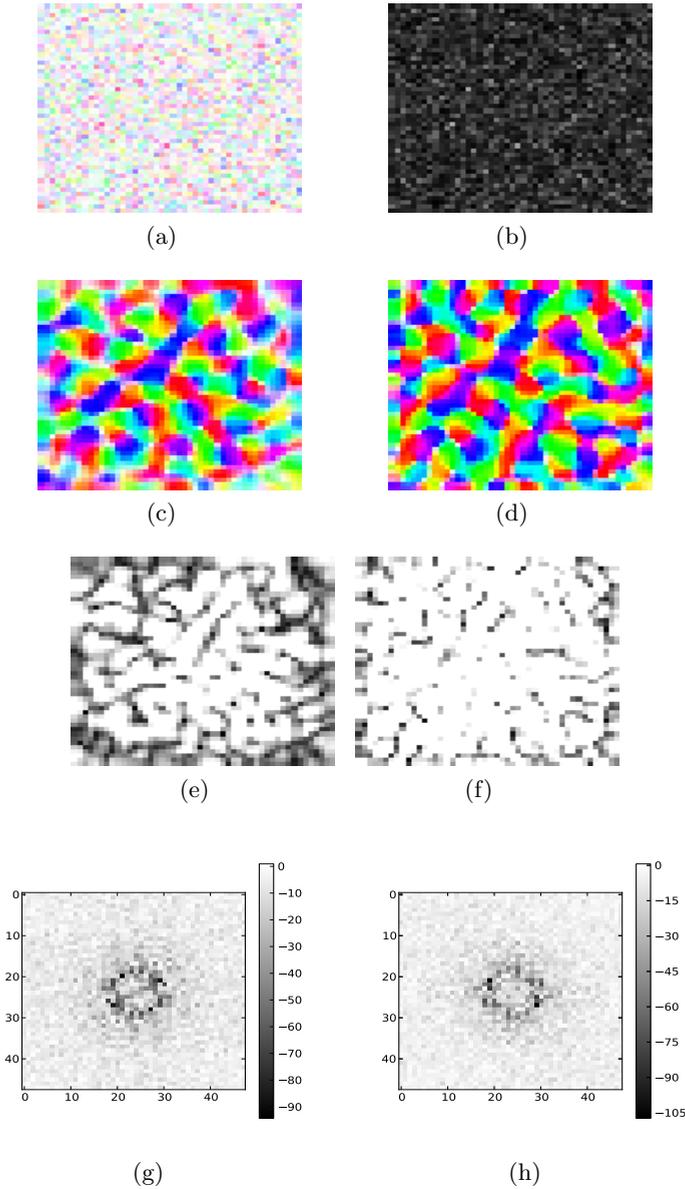


Fig. 4. Orientation Preference/Selectivity maps and Fourier spectrums. (a) initial orientation preference/selectivity map before the training; (b) initial selectivity map; orientation preference/selectivity maps after the training of the architecture without the LGN level (c) and with the LGN level (d); selectivity maps after the training of the architectures without the LGN level and (e) with the LGN level (f); Fourier spectrums of the orientation preference/selectivity maps for the architecture without the LGN level (g) and with the LGN level (h). The selectivity maps show that before training (b) the neurons are unselective (i.e. dark), and after training (e) and (f) nearly all of the neurons are highly selective (light)

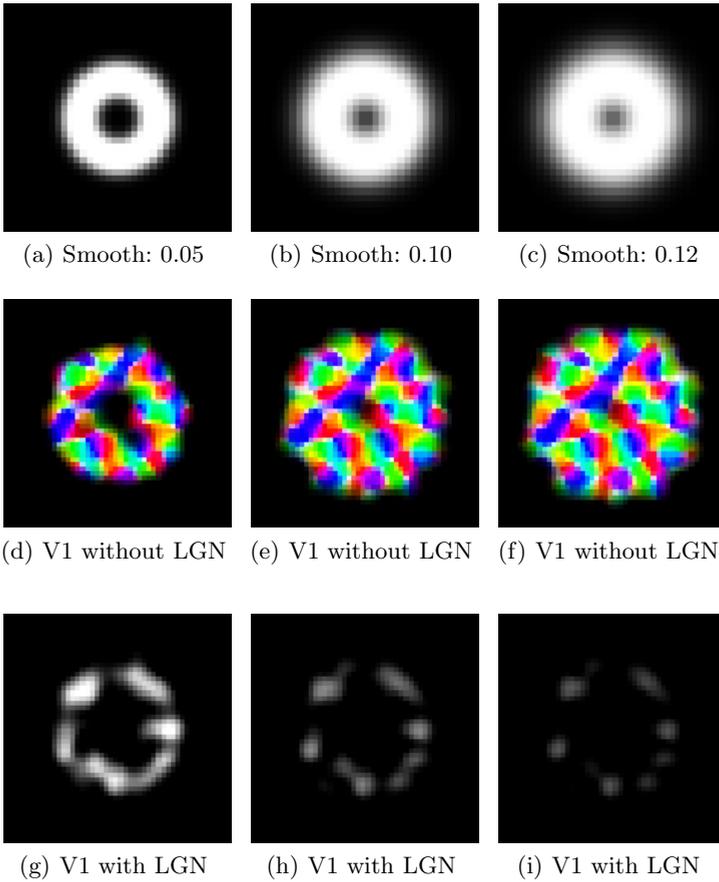


Fig. 5. V1 activity testing with the same ring shaped pattern, increasing its level of smoothing. Ring test patterns corresponding to (a) smoothing of 0.05, (b) smoothing of 0.10, (c) smoothing of 0.12. The middle row of the figure show the V1 activity responses corresponding to the architecture without the LGN level. The last row of the figure show the V1 activity responses corresponding to the architecture with the LGN level. The V1 answer is patchy due to the orientation preference and selectivity of the V1 neurons.

6 Discussion

The orientation preference/selectivity maps of Fig. 4 show that the training was successful in obtaining smoothly varying orientation and selectivity maps. An interesting property of orientation maps measured in animals is that their Fourier spectrums usually show a ring shape, because the orientations repeat at a constant spatial frequency in all directions. The subfigures 4(g) and 4(h) show ring shaped Fourier spectrums for both architectures, which thus indicates

that the orientation preference/selectivity maps are similar to biological maps. The selectivity map for the neurons in the architecture with the LGN level (Fig. 4(f)), is lighter than the selectivity map for the architecture without the LGN level (Fig. 4(e)). This result indicates higher selectivity for the neurons in the architecture with the LGN level, which coincides with previous results in [1].

The V1 activities of Fig. 5, show that when the level of smoothing is increased, the recognition of the ring shaped pattern is more difficult for both architectures. The architecture without the LGN level gets the worst performance, because when the level of smoothing is increased, the answer of the V1 seems a disk instead of a ring. Thus, the architecture with the LGN level seems more robust to gradual changes in illumination in a ring shaped pattern. This result gives more strength to the result in [1] regarding the robustness of an architecture with LGN in natural environments, because smoothing is an important visual characteristic in natural images.

7 Conclusions and Future Work

In this article, the performance of hierarchical architectures for computer vision based on LISSOM was analyzed. In particular, the importance of the LGN level for the robustness under gradual changes in illumination was analyzed. Changes in illumination are common in everyday environments and are therefore a crucial factor that affects the recognition of objects in everyday environments. The LISSOM architecture with the LGN level showed a good level of robustness for ring shaped pattern recognition under changes of illumination.

Only single ON and OFF channels are used in LISSOM simulations in this article. Multiple channels representing different spatial frequencies are currently being analyzed to be included in the models, in order to improve the computational early visual processing.

Regarding the velocity of the answer of V1, it was very fast, after the training, which could be suitable for real time applications. However, the dimension of the networks was small, and the dimension of the network affects the capacity of the model to recognize patterns. Thus, parallel implementation of the model are currently being analyzed to increase its velocity and use larger networks in realistic environments.

Processing in the brain in general and visual processing in particular is organized in a hierarchical fashion, from simple localized features to complex, large scale features. The visual system consists of a hierarchy, in which neurons in early visual areas extract simple image features (e.g., orientation) over a small local region of visual space, which are then transmitted to neurons in higher visual areas responding to more complex features (e.g. shape) over a larger region of visual space. Hierarchical representations can derive and organize features at multiple levels. However, although all neurophysiologic evidence suggests that in the human visual system quite a number of levels are realized, it has turned out that the design and/or learning of such deep hierarchical systems is a very difficult task. Most existing computer vision systems are 'flat' (e.g., having rather

simple features as input) and hence cannot make use of the advantages connected to deep hierarchies. Here in particular the generalization capabilities are crucial for any form of cognitive intelligence. As a consequence, the issue of establishing deep hierarchies is considered as one major challenge for the establishment of truly cognitive systems. This paper deals with the computational modeling at the level of the early visual processing, and is at the base of a hierarchical cognitive system. Thus, the study, design and optimization of early visual processing levels is also relevant to achieve truly cognitive systems.

References

1. Miikkulainen, R., Bednar, J., Choe, Y., Sirosh, J.: *Computational Maps in the Visual Cortex*. Springer (2005)
2. Ringach, D.: Mapping receptive fields in primary visual cortex. *The Journal of Physiology* (558), 717–728 (2004)
3. Miikkulainen, R.: Self-organizing process based on lateral inhibition and synaptic resource redistribution. In: Kohonen, T., Makisara, K., Simula, O., Kangas, J. (eds.) *Proceedings of the 1991 International Conference on Artificial Neural Networks* (1991)
4. Miikkulainen, R., Bednar, J., Choe, Y., Sirosh, J.: Self-organization, plasticity, and low-level visual phenomena in a laterally connected map model of the primary visual cortex. *Psychology of Learning and Motivation* 36, 257–308 (1997)
5. Sirosh, J.: A self-organizing neural network model of the primary visual cortex. PhD thesis, Department of Computer Sciences, The University of Texas at Austin, Austin, TX (1995)
6. <http://www.topographica.org>