A Model of Angle Selectivity in Area V2 with Local Divisive Normalization

Mohsen Malmir, Saeed Shiry Ghidary

Abstract— Efficient coding hypothesis states that the goal of sensory system of the brain is to remove redundancies in the sensory input. Several models tried to remove redundancy in the visual input and successfully modeled the functional properties of neurons in the primary visual cortex. However there has been no progress to extend these models to simulate the properties of neurons in extrastriate visual areas. In this paper, we propose that visual cortex tries to remove higher order dependencies in a hierarchical architecture. In each layer a nonlinear mechanism removes redundancies in a local neighborhood. We used the biologically plausible divisive normalization mechanism in a two layer model network to remove nonlinear dependencies in the input. Units in this model can simulate the responses of neurons in area V2 to angle stimuli.

I. INTRODUCTION

VENTRAL visual pathway is thought to be the locus for processing object identity information in the visual cortex [1]. This processing pathway extends from V1 in the occipital cortex through V2 and V4 to the Inferotemporal cortex [2]. Along this pathway, receptive fields of neurons grow in size and the preferred stimuli changes from simple edge segments to complex shapes like faces [3]–[6]. These observations have led to the idea of hierarchical visual processing in the object recognition pathway [1]–[3], [7], [8].

Information theoretic methods are an approach to study the functions of visual system (and other sensory systems). These methods are usually based on the Barlow’s “efficient coding” hypothesis [9] which itself was motivated by the information theory. In his hypothesis, Barlow stated that sensory input to the brain is largely redundant because of the structural regularities in the natural environment. This redundant input may place a burden on the limited processing capabilities of the brain. One way for brain to overcome this problem is to remove these redundancies in the sensory system and feed only the most informative features to the higher levels. This theory has been described in various forms by others [10]–[13].

There are a number of successful applications of the “efficient coding” hypothesis which have modeled the functional properties of neurons in the primary visual cortex. In an early study, Olshausen and Field tried to represent the input image based on sparse activation of linear filters [14]. They found localized band pass basis functions which were very similar to the selectivity of simple cells in the primary visual cortex. Similar results were achieved in a work by Bell and Sejnowski who linearly transformed the input image to a set of independent components [15].

Hyvärinen and Hoyer used energies of responses of filters in independent subspaces to model the shift and phase invariance properties of complex cells [16]. Schwartz and Simoncelli tried to model nonlinear dependencies in natural images using a nonlinear behavior observed in the neurons of visual and auditory systems [17]. They used divisive normalization to remove variance dependency among responses of a set of linear filters. Their model could account for a number of nonlinear physiological behaviours that had been observed in primary visual and auditory cortices.

Despite great success in modeling different properties of neurons in the primary visual cortex, there has been no progress to extend these results to the properties of neurons in extrastriate visual cortical areas. Some studies tried to learn higher order structures in natural images through capturing variance dependencies among responses of linear filters [18]–[20]. However there were no similarities between the learned features in these models and the properties of neurons in extrastriate visual areas. In order to model the properties of neurons in higher order areas of visual cortex, we develop a model to remove higher order dependencies in the input image. We use divisive normalization in a hierarchical model which allows removing higher order dependencies in the successive stages of the hierarchy.

II. HIGH ORDER DEPENDENCIES AND HIERARCHICAL DIVISIVE NORMALIZATION MODEL

The problem with linear redundancy reduction methods like ICA is that, they cannot yield independent components with a simple linear transformation. Therefore some residual redundancies remain between responses of linear filter. Several methods have used nonlinear interactions to remove these residual redundancies to yield independent components [17], [18], [21], [22]. These models assume that the responses of distant linear filters are independent but the responses of adjacent filters exhibit variance dependency.

The question is whether the responses of distant filters are really independent or they have some form of nonlinear...
dependency that is not easily detectable. In order to answer this question, we implemented a local divisive normalization mechanism as described in [17]. We calculated the joint histogram of the activities of two filters before and after normalization and found that before normalization the responses of linear filters seemed to be independent (Fig. 1a). However as can be seen in the joint histogram of normalized responses in Fig. 1b, variance of the normalized responses of one filter depends on the normalized responses of the other.

An intuitive explanation for this observation is that there is a higher order dependency between linear filters responses beyond variance dependency. This higher order dependency cannot be detected in the joint histogram of the responses of two linear filters because it does not affect the mean or variance of these responses. In the topographic ICA or divisive normalization model it was assumed that ICA can result in independency between responses of these filters. Yet we can see that there is a nonlinear dependency between energies of these responses which cannot be eliminated using these transformations. This dependency which was hidden in the joint histogram of linear filters responses is revealed in the joint histogram of energies of them.

![Fig. 1. Joint histogram of responses of two distant filters (a) before normalization and (b) after normalization](image)

In order to remove this redundancy we added a second layer of divisive normalization to the model of Schwartz and Simoncelli. Normalized responses of linear filters in the first layer are fed to the units in the second layer. Divisive normalization is used to remove the nonlinear variance dependency between activities of units in the second layer. Such an application of divisive normalization has the ability to remove high nonlinear dependencies among responses of linear filters. The advantage of using this method is that it can remove higher nonlinear dependencies with a biologically plausible mechanism.

III. MODEL DESCRIPTION

The model we propose here is an extension to the model of Schwartz and Simoncelli. In this model the input image is transformed with a bank of linear filters similar to those found in ICA [15]. Then responses of these linear filters are transformed with a nonlinear divisive normalization mechanism in order to remove the variance dependency between them. The variance of responses of each linear filter is modeled as a sum of squared responses of filters in a neighboring region as in (1).

$$\text{var}(L_x | L_y, \bar{y} \in C_x) = \sum_{y \in C_y} w_{xy}^2 L_y^2 + \sigma_x^2$$

(1)

where $L_x$ and $L_y$ are the responses of linear filters $x$ and $y$, $C_y$ is a neighboring area around filter $x$, $w_{xy}$ is the weight of response of filter $y$ in normalizing responses of $x$, and $\sigma_x^2$ is a part of the variance of $x$ which is independent of the other filters responses. In order to remove the variance dependency, responses of filters are transformed using (2).

$$R_x = \frac{L_x^2}{\sum_{y \in C_y} w_{xy}^2 L_y^2 + \sigma_x^2}$$

(2)

Parameters of this transformation are calculated using maximum likelihood estimate over an ensemble of natural images as in (3).

$$\hat{w}_{xy}, \hat{\sigma}_x = \arg\max_{w_{xy}, \sigma_x} \prod_i P(L_x(i) | L_y(i), \bar{y} \in C_x)$$

(3)

where $P$ denotes the conditional probability distribution function of linear filters responses. For simplicity of computations the joint distribution function in the first and second layer is set to the normal distribution.

Normalization in the second layer of the model incorporates signals from larger vicinities compared to the first layer. Notice that simply extending the normalization area in the first layer does not remove this higher order dependency because it involves the variance of energies of linear filters not the variance of responses themselves. In the second layer responses are normalized using (4).

$$I_{\eta} = \frac{R_{\eta}^2}{\sum_{y \in C_y} w_{xy}^2 R_y^2 + \sigma_\eta^2}$$

(4)

where $I_{\eta}$ is the normalized responses of filter $\eta$ after normalization and $C_{\eta}$ denotes the region in the second layer in which responses are used to normalize the responses of $\eta$.

We calculated the parameters of normalization in the first layer by likelihood estimation over an ensemble of natural images. After learning these parameters, the responses of filters in the first layer were calculated and used to estimate the parameters of the second layer. Estimation of parameters of normalization in the second layer was done using
likelihood estimate over the responses of the first layer.

IV. COMPARISON OF MODEL UNITS WITH REAL V2 NEURONS

We compared the properties of the units in the second layer of the proposed model to the properties of neurons in the area V2. This is an important area in the visual cortex and is as large as area V1. However there are few studies which have systematically examined the selectivity of neurons of V2. The problem in the study of stimulus selectivity in this area is to select a stimulus set which can elicit neuronal firings. It has been suggested that this area might be selective to complex stimuli such as gratings and curves [4], [23]–[25].

One of the most prominent studies of visual area V2 is the work done by Ito and Komatsu who have examined the selectivity of neurons in this area to angle stimuli [26]. They used a set of angles with different orientations and angle widths to find if neurons in V2 are selective to angles. They found that most of the V2 neurons were responsive to these stimuli and some neurons demonstrated sharp selectivity to a specific angle stimulus. We adapted their stimulus set (Fig. 2) to examine the selectivity of units in the second layer of our proposed model. This stimulus set is composed of 66 angles that have two component half lines. These half lines vary from 0˚ to 330˚ with 30˚ steps. Angles in the upper triangle in Fig. 2 are replicated in the lower triangle for better visualization.

Fig. 2. Stimuli set used to examine the selectivity of model V2 neurons. Angles in the shaded region are a mirror image of angles in the white region.

Fig. 3a shows the responses of a typical unit in the second layer of proposed model to the stimuli set. This unit is indeed selective to a 120˚ angle composed of 150˚ and 270˚ half lines. The topmost row and leftmost column in Fig. 3 display the responses of this unit to the individual half lines. This unit showed no selectivity to the 270˚ and a mild response to the 150˚ half line. In this case, the unit is selective to the combination of these two half lines. The response of this unit to the optimal stimulus is a nonlinear function of the responses to the constituent half lines. A similar nonlinear behavior was observed in the responses of neurons in area V2.

Fig. 3b displays the responses of another model unit to the stimulus set. The maximum response in this unit is elicited by an angle composed of 0˚ and 330˚ half lines. This unit also responded to other angles which have a 0˚ half line component. This is evident in the Fig. 3b that the responses are elongated along the row and column corresponding to the 0˚ half line. This means that this unit is selective to a 0˚ half line and its response to the optimal angle is a nonlinear function of the constituent half lines. A similar behavior was found and reported in [26].

A. Distribution of Optima Angle Stimuli in the Model

We compared the distribution of optimal angle width over a sample population of model units with those observed in real V2 neurons reported in [26]. Fig. 4 shows this distribution for model units and real V2 neurons. A large number of model units (41%) showed maximum response to 30˚ degree angle. This is close to the 43% of real V2 neurons. Wide angles between 60˚ and 150˚ were optimal stimulus in 49% of model units and 41% in real V2 neurons. Long bars were optimal stimulus for only 9% of model units and 16% of real V2 neurons.

Fig. 4. Comparison of optimal angle stimuli distribution in (a) real V2 neurons regenerated from [26] and (b) in model units.
are close to the data acquired in [26]. They reported that 21% of peak response areas include less than 4 angles.

V. CONCLUSION

“Efficient coding” hypothesis has greatly influenced the direction of research in the study of brain sensory system. Several studies used this hypothesis and successfully modeled a number of properties of neurons in the visual cortex. Primary works used linear transformation to remove redundancies in the input image [14], [15]. Nonlinear behavior of neurons in the visual cortex inspired subsequent models to use nonlinear interactions to remove redundancy. These models successfully generated nonlinear units which replicated the behavior of neurons in the primary visual cortex [16], [17].

Dependencies in the natural environment are highly nonlinear and we cannot claim that one stage of nonlinear transformation can completely remove them. A possible mechanism in the brain to remove these dependencies is a hierarchical processing system in which successive layers remove higher order dependencies in the input. In each layer nonlinear interactions such as divisive normalization are used to remove the visible redundancies in the input. Moreover these nonlinear interactions reveal higher order redundancies for the next layer. In this way a simple nonlinear interaction can remove higher order redundancies.

We implemented a two layer neural network and used divisive normalization in each layer to remove variance dependencies between network unit’s responses. Output of the first layer is calculated by normalizing the responses of a set of linear filters in a neighboring region. In the second layer these normalized responses are again normalized in larger vicinities. This mechanism can remove dependencies between variances of energies of linear filters responses. We compared the properties of units in the second layer of proposed network to the properties of real V2 neurons and found that these units could simulate a number of behaviors of real V2 neurons.

One important property observed in neuronal responses of area V2 is that responses to angle stimuli are a nonlinear function of responses to the constituent half lines. Units in the second layer of the proposed model can simulate similar nonlinear behaviours. A large number of V2 neurons are selective to sharp angle stimuli. This was also observed in the selectivity of units in the model. Distribution of the size of peak response area was similar in the model units and real V2 neurons. These results indicate that our proposed model can simulate the properties of neurons in area V2.

REFERENCES


