Chapter V
A Selective Sparse Coding Model with Embedded Attention Mechanism

Qingyong Li
Beijing Jiaotong University, China

Zhiping Shi
Chinese Academy of Sciences, China

Zhongzhi Shi
Chinese Academy of Sciences, China

ABSTRACT
Sparse coding theory demonstrates that the neurons in the primary visual cortex form a sparse representation of natural scenes in the viewpoint of statistics, but a typical scene contains many different patterns (corresponding to neurons in cortex) competing for neural representation because of the limited processing capacity of the visual system. We propose an attention-guided sparse coding model. This model includes two modules: the non-uniform sampling module simulating the process of retina and a data-driven attention module based on the response saliency. Our experiment results show that the model notably decreases the number of coefficients which may be activated, and retains the main vision information at the same time. It provides a way to improve the coding efficiency for sparse coding model and to achieve good performance in both population sparseness and lifetime sparseness.

INTRODUCTION
Understanding and modeling the functions of the neurons and neural systems are one of the primary goals of cognitive informatics (CI) (Wang 2002, 2007; Wang and Kinsner 2006). The computational capabilities and limitations of neurons, and the environment in which the organism lives are two fundamental components driving the evolution and development of such systems. The researchers have broadly investigated them.
The utilization of environmental constraints is most clearly evident in sensory systems, where it has long been assumed that neurons are adapted to the signals to which they are exposed (Simoncelli 2001). Because not all signals are equally like each other, it is natural to assume that perceptual systems should be able to best process those signals that occur most frequently. Thus, it is the statistical properties of the environment that are relevant for sensory process of vision perception (Field 1987; Simoncelli 2003).

Efficient coding hypothesis (Barlow 1961) provides a quantitative relationship between environmental statistics and neural processing. Barlow for the first time hypothesized that the role of early sensory neurons was to remove statistical redundancy in the sensory input. Then, Olshausen and Field put forward a model, called sparse coding, which made the variables (equivalence of neurons stimulated by the same stimulus in the neurobiology) be activated (i.e., significantly non-zero) only rarely (Olshausen 1996). This model is named SC here. Vinje’s results validated the sparse properties of neural responses under natural stimuli conditions (Vinje 2000). Afterwards, Bell brought forward another sparse coding model based on statistical independence (called SCI) and obtained the same results as Olshausen and Field’s model (Bell 1997). More recent studies can be seen in survey (Simoncelli 2003).

However, Willmore and Tolhurst (Willmore 2001) argued that there were two different ways for 'sparseness': population sparseness and lifetime sparseness. Population sparseness describes codes in which few neurons are active at any time and it is utilized in Olshausen and Field’s sparse coding model (Olshausen 1996); while lifetime sparseness describes codes in which each neuron's lifetime response distribution has high kurtosis, which is the main contribution in Bell’s sparse coding model (Bell 1997). In addition, it is proved that lifetime sparseness was uncorrelated with population sparseness. Just as figure 3.a shows the number of variables, which have large values produced by sparse coding model and are possible to be activated, is relatively large compared with the computation capacity of neurons. Though, the kurtosis of every response coefficient is also high. So, how to reduce both population sparseness and lifetime sparseness at the same time to retain the important information as much as possible is a valuable problem in practice.

Vision attention mechanism is an active strategy in information processing procedure of brain, which has many interesting characteristics such as selectivity and competition. Attention is everywhere in the visual pathway (Britten 1996). Furthermore, a typical scene within the neuron’s classic receptive field (CRF) contains many different patterns that compete for neural representation because of the limited processing capacity of neurons in the visual system. So, integrating attention mechanism into sparse coding framework to reduce the population sparseness and to improve the coding efficiency sounds reasonable and essential.

In this chapter, we extend sparse coding principle combining the vision attention. We first model the sampling mechanism of retina by a non-uniform sampling module; then, we implement a bottom-up attention mechanism based on the response saliency of the sparse coefficient. The diagram is illustrated in figure 1. This model has two main contributions:

1. Modeling the vision attention in the framework of sparse coding.
2. Improving the population sparseness of the response coefficient in the same time retaining the most efficient information.

Figure 1. The diagram of the model