Available online at www.sciencedirect.com

SCIENCE DIRECT®

Neurocomputing 69 (2006) 1219–1223

NEUROCOMPUTING

www.elsevier.com/locate/neucom

Storing and restoring visual input with collaborative rank coding and associative memory

Martin Rehn[a,*], Friedrich T. Sommer[b]

[a]*KTH, NADA, Royal Institute of Technology, SE-100 44 Stockholm, Sweden*
[b]*Redwood Center for Theoretical Neuroscience, University of California, Berkeley 132 Barker, Berkeley, CA 94720-3190, USA*

Abstract

Associative memory in cortical circuits has been held as a major mechanism for content-addressable memory. Hebbian synapses implement associative memory efficiently when storing sparse binary activity patterns. However, in models of sensory processing, representations are graded and not binary. Thus, it has been an unresolved question how sensory computation could exploit cortical associative memory.

Here we propose a way how sensory processing could benefit from memory in cortical circuitry. We describe a new collaborative method of rank coding for converting graded stimuli, such as natural images, into sequences of synchronous spike volleys. Such sequences of sparse binary patterns can be efficiently processed in associative memory of the Willshaw type. We evaluate storage capacity and noise tolerance of the proposed system and demonstrate its use in cleanup and fill-in for noisy or occluded visual input.

Keywords: Sensory coding; Attractor memory; Rank coding; Sequence memory; Data compression

---

1. Introduction

The microcircuitry of the cerebral cortex shows extensive recurrent connectivity between pyramidal cells in layer II/III. These connections are plastic and have been shown to obey a temporally asymmetric Hebbian learning rule [2]. Associative memories are computational models that describe how cortical circuits may exploit these plastic connections to perform memory operations. Recently, physiological experiments in slice have lent support to these models by showing that activity organizes in repeatable sequences of activity patterns [7]. Under certain conditions, neural associative memories can efficiently store and retrieve large numbers of patterns of neural activity. These conditions are in particular that the patterns are binary and sparse, i.e., that the stored patterns share a low ratio of active cells [17,10,11]. At first glance these conditions seem ill-suited for sensory processing. Visual input, for example, when represented by Gabor-type filters, is graded and nonsparse. However, the recent discovery of discrete so-called UP states in visual cortex provides at least indirect evidence that binary and sparse information processing could be relevant even in early sensory processing [4]. To date, no biologically plausible memory model has been proposed that can store large numbers of chunks of analog raw sensory data, such as images. The aim of this paper is to propose such a model.

It has been shown that principles of efficient coding [14] and also faster but suboptimal techniques of signal representation, such as matching pursuit can sparsify sensory neural representations. These mechanisms can be neuronally implemented by lateral inhibition. Matching pursuit, in particular, has been suggested to convert sensory input into temporal sequences of spikes [13]. For efficient processing of sensory information we propose a combination of a new model of sparse visual coding, extending the model based on matching pursuit, and sparse sequence associative memory [1,18].

---

*Corresponding author. +46 7 90 7784; fax: +46 8 790 0930.
E-mail addresses: rehn@nada.kth.se (M. Rehn), fsommer@berkeley.edu (F.T. Sommer).

0925-2312/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2005.12.080

2. Image coding

2.1. Matching pursuit for the spike coding of visual input

Perrinet et al. [13] have proposed a model of the visual cortex based on a method of signal representation known as matching pursuit. In this model, neurons spike one at a time, each spike being elicited from the neuron that is most strongly excited by the sensory input. Neurons that spike inhibit other cells with similar receptive fields ("explaining away"). The representation of visual input by matching pursuit can be written as

$$\mathbf{x} = \sum_{n=0}^{m-1} \langle \mathbf{r}_n, \mathbf{\Psi}_{\gamma_n} \rangle \mathbf{\Psi}_{\gamma_n} + \mathbf{r}_m, \tag{1}$$

$$\gamma_n = \mathrm{argmax}_{\gamma_n} \langle \mathbf{r}_n, \mathbf{\Psi}_{\gamma_n} \rangle, \tag{2}$$

where $\mathbf{\Psi}$ is a set of basis functions, $\mathbf{x}$ is the visual input and $\mathbf{r}_n$ is the residual after the n:th spike. Eq. (2) indicates that the neuron γ_n to spike next will be the one corresponding to the basis function most similar to the current residual $\mathbf{r}_n$. To determine how many spikes should be used in the sequence, one can optimize the representation based on a cost function. We use the function

$$E(\mathbf{x}, \mathbf{b}) = \frac{1}{2} \sum_{i=1}^{m} \left(x_i - \sum_{j=1}^{n} b_j \Psi_{ji} \right)^2 + f(\mathbf{b}), \tag{3}$$

where the first term quantifies the quality of the representation $\mathbf{b}$, as generated by matching pursuit; $b_j = \sum_{n:\gamma_n = j} \langle \mathbf{r}_n, \mathbf{\Psi}_{\gamma_n} \rangle$. The second term quantifies the metabolic cost of the representation. Assuming that each spike is associated with a fixed metabolic energy cost, we just use the spike count

$$f(\mathbf{b}) = \theta \|\mathbf{b}\|_{L0}. \tag{4}$$

Depending on the choice of the sparseness parameter θ this coding scheme can produce sparse codes. The codes have few nonzero elements but are not yet binary. They contain either zeros or analog expansion coefficients $\mathbf{b}$. It has been demonstrated, however, that the exact analog values are not needed for faithful reconstruction of visual input. They may be replaced by mean values from a rank ordered histogram of coefficients that is averaged over many visual inputs [12]. Thus, an input pattern in a patch of the visual field can be coded by a temporal sequence of spikes, where only the spike order is significant.

2.2. Collaborative rank coding of image fragments

To use spike coding based on matching pursuit in combination with efficient associative memory, the coding strategy explained so far has to be extended. In our model of *collaborative* rank coding, cortical regions processing different patches of visual input collaborate to form spatio-temporal patterns. In our model we assume that an image is tiled by small nonoverlapping patches in the visual space that are processed in parallel by sets of neurons in different regions of primary visual cortex. For each patch, matching pursuit is used to determine the next spiking neuron. But spike timing in different patches is not independent as in the Perrinet model. We include a global, synchronizing influence which could be realized in the cortex by local thresholds that are synchronously oscillating [3]. The effect of this collaborative rank coding is that spikes in different patches organize into synchronized volleys of spikes. The number of spikes per volley can be regulated by the degree of threshold modulation. In our computer model we simply group the k largest coefficients in the first volley, the following k largest in the second and so on until less than k nonzero coefficients remain, at which point the sequence is truncated. For image number μ, we define r^{μ} as the index vector of the nonzero coefficients in $\mathbf{b}$, ordered according to descending magnitude. We denote by $|r^{\mu}|$ the cardinality of nonzero coefficients. The input image $\mathbf{x}^{\mu}$ is then represented by a sequence of patterns $\xi^{\mu}(t)$ of length T^{μ}; $t \in [1 \dots T^{\mu}]$:

$$\xi_i^{\mu}(t) = \begin{cases} 1 & r^{\mu}(i) \in ((t-1)k,\, tk], \\ 0 & \text{otherwise}, \end{cases} \tag{5}$$

$$T^{\mu} = \lfloor |r^{\mu}|/k \rfloor. \tag{6}$$

This k-winner-take-all coding strategy implements the collaborative rank coding.

For reconstruction we form a rank code lookup table for the analog coefficients, based on the volley index (see Fig. 1). Comparing the error bars of collaborative and non-collaborative rank coding in Fig. 1 the collaboration seems to increase the accuracy of the lookup. An analysis of the rank code statistics of these coding schemes will be given elsewhere. Most importantly, the collaborative rank coding allows for the reconstruction of a stimulus, given only an joint spike volley sequence and the collaborative rank code lookup table.

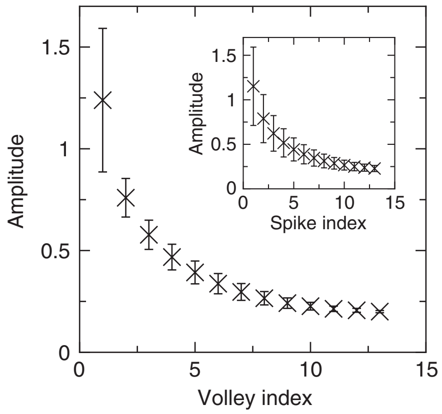

Fig. 1. Coefficient amplitudes as a function of volley index. Error bars show one standard deviation. The inset figure is for non-collaborative rank coding. Note the larger variance for the latter.