Independent component analysis of temporal sequences forms place cells

András Lőrincz*a,*, Gábor Szirtesa, Bálint Takácsa, György Buzsákiib

aDepartment of Information Systems, Eötvös Loránd University, Pázmány Péter sétány 1/D, Budapest, H-1117, Hungary
bCenter for Molecular and Behavioral Neuroscience, Rutgers, The State University of New Jersey, 197 University Ave., Newark, NJ 07102, USA

Abstract

It has been suggested that sensory information processing makes use of a factorial code. It has been shown that the major components of the hippocampal-entorhinal loop can be derived by conjecturing that the task of this loop is forming and encoding independent components (ICs), one type of factorial codes. However, continuously changing environment poses additional requirements on the coding that can be (partially) satisfied by extending the analysis to the temporal domain and performing IC analysis on concatenated inputs of time slices. We use computer simulations to decide whether IC analysis on temporal sequences can produce place fields in labyrinths or not. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Hippocampus; Cognitive map; Generative network; Independent component analysis; Temporal sequences

1. Introduction

It has been argued (see, e.g., [1,5], and references therein) that the processing of sensory information makes use of a factorial code for filtering. Such filters would code, e.g., the properties of natural scene ensembles that occur independently of each other. The argument is supported by recent numerical studies [6,16,2,17]. It was found that training filters on natural images by using the related techniques of sparse representation formation and independent component (IC) analysis (ICA) produces receptive
fields alike to the receptive fields of the primary visual cortex. However, if the analysis is to be implemented by neural networks, a new condition arises, namely, inputs to the analyzing network should be drawn independently. This is not fulfilled in a continuously changing environment. In other words, if temporal correlation between the inputs exists then the inputs should be shuffled before learning. Shuffling requires the storage of all of the inputs before learning could start. This condition means batch learning and contradicts on-line learning. If inputs to the analyzing network are provided in discrete time instances then ICA can be extended to temporal sequences by concatenating a number of subsequent inputs. Concatenation will be called embedding and the number of concatenated inputs as embedding dimension. ICA performed in this larger dimensional space can project onto the original (non-concatenated) number of dimensions simultaneously. This extension of ICA to temporal sequences decreases temporal correlation between the components of the representation within the time window of embedding. IC analysis on temporal sequences may serve as the basis of a hierarchical IC procedure that creates factorial code and removes temporal correlations in a systematic fashion. Interestingly, Hateren and Ruderman found [7] that extending IC analysis to temporal sequences improves the agreement between the emerging receptive fields and the experimentally found receptive fields of the primary visual cortex.

It has been shown that the major components of the hippocampal-entorhinal loop can be derived by conjecturing that the forming and the encoding of ICs are the tasks of this loop [14,15]. According to the model, the hippocampal-entorhinal loop forms a reconstructing dynamical network that makes use of orthogonal projections for whitening and separation, and computes a top-down generative IC code, that forms long-term memories. Here we examine one aspect of this model by asking the question whether IC analysis on temporal sequences can produce place fields.

2. Methods

A random model labyrinth was generated with a circular path. We have assumed that sensory information at time t is made of local observations. At each point of the labyrinth, the processed sensory information (PSI) consists of eight bits. The first four bits indicate the presence or the absence of the walls at the north, east, south and west side of the ‘animal’. The presence (absence) of a wall makes the respective input 1 (0). The next four bits indicate the direction (north, east, south or west) of the path. The resulting eight bit inputs are convolved with a 0.5, 1.0, 0.5 filter along the path that decreases the precision of local information. In the computations, the FastICA algorithm [9] was utilized. Concatenated temporal sequences of PSIs were input to the algorithm. Embedding depths varied from 1 (8 dimensional input) to 24 (24 x 8 = 192 dimensional input). The IC transformation matrix was trained on the labyrinth of Fig. 1(a). Input concatenation corresponded to one (counter-clockwise, referred to as forward) direction. After training, the generated matrix was used for further analysis. A normalization procedure was used: The absolute values of the IC outputs were added at time t and, in turn, each output was divided by this sum at each
Fig. 1. (a) The arrows show the random path used. Processed sensory information is made of two parts. The first four digits show the presence or the absence of the walls to the south, east, north and west side of the ‘animal’. The presence (absence) of a wall makes the respective input 1 (0) (see second inlet). The second four digits show the direction of the path (south, east, north or west) (third inlet). The length of the path is 58 steps. (b) Histograms of IC output. Upper left: trained (original) direction, upper right: novel (reverse) direction for embedding depths 8, 16, and 24. Lower left: exponential fits to curves of the upper left figure.

time instant. We assumed that such normalization is achieved, e.g., by inhibition [3]. IC outputs can be either positive or negative, which has many possible neural interpretations. Olshausen and Field [17,16] say that deviation from the average firing rate provides the sign of the signal. Another interpretation can be provided by positive coding networks [4]. That is why only the absolute values of the IC outputs are shown in our figures. Testing was performed by employing the trained ICA matrix on several different labyrinths. The firing rate distributions in novel labyrinths were the same (within statistical error) then the firing rate distribution in the original labyrinth with concatenation in reverse (clockwise, i.e., reverse) direction. Here, inputs to the IC matrix in reverse direction will be referred to as ‘novel’ inputs.

3. Results

The first experiment was concerned with the forming of place cells in the case when the training inputs were generated both in forward and in reverse directions. Most ICs formed localized ‘place cells’ in the trained direction, which means their activity was high at around a given point along the path. Alas, there was no sign change in and around this point that supports the idea of positive coding [4]. In reverse direction, none of the ICs can be termed ‘place cell’. In turn, these ‘receptive fields’ were direction sensitive. That is, if an IC was strongly responding in one direction, then it was much less responsive in the other (see also Fig. 2).
The IC output distributions were analyzed. It was found that in the training direction the IC outputs follow exponential distribution, whereas in the reverse direction, the distributions are peaked at non-zero values and resemble to truncated Gaussian distributions (Fig. 1(b)).

Fig. 2 depicts the absolute values of the IC outputs normalized when using 24 embedding dimension. We found that larger embedding depth improved place cell formation. This was not a monotone feature however. Place cell formation can fully disappear when embedding time depth corresponds to the round trip time of the labyrinth.

4. Discussion

Our computer simulations demonstrate that place cells can be formed by IC analysis of temporal sequences. These cells are formed according to the particular form of training. In a one-dimensional labyrinth (‘single path’) ICs represent places and directions. IC activities are peaked in localized regions and the activities are conditioned on the direction of motion. Indeed, this is the case for the radial arm maze, linear track and running wheel, i.e., when the animal traverses repeated paths. We have studied this case here. Directed place cells were formed also when both directions were used during the training session. The formation of invariant place fields that can develop in other scenes, e.g., in free two-dimensional areas is highly non-trivial if only relative directions and relative position differences are provided for IC analysis. One important issue is whether invariant place cells can be formed by IC analysis alone when both local and global information are input to the algorithm. This point remains to be seen.
Hippocampal principal cells may be able to analyze temporal sequences and thus these cells may be able to perform IC analysis on temporal sequences. Distance on the dendritic trees of principal cells measured from the soma in the CA3 and CA1 subfields can be viewed as time on the 10–100 ms time scale because of propagation delays along the dendritic trees [8]. Recent modeling efforts of Jaffe and Carnevale [10] indicate that dendritic signals far from the soma are not damped relative to dendritic signals close to the soma in principal cells of CA3. Arguments that propagation delays, in effect, make the dendritic tree to concatenate information from different time depths have also been published [13,11,12].

We found that in trained cases the distribution of IC component amplitudes follows exponential behavior whereas in novel situations the distribution is peaked and is more reminiscent to a truncated Gaussian distribution. In the brain, for certain sets of inputs and for certain neurons, firing rate distributions resemble better to exponential distributions, whereas in other cases these resemble better to a truncated Gaussian distributions [18]. Our findings suggest that the novelty content of the information may play an important role in the firing rate distribution. In this respect it is important that neither exponential nor Gaussian distribution can be fitted to most of the experimental data [18].

References


Andras Lorincz is at Eotvos Lorand University, Budapest. His background is experimental and theoretical physics and physical chemistry. He has received his Ph.D. from Eotvos Lorand university in physics in 1978. He has been involved in artificial intelligence research for about 10 years. He is presently leading the Neural Information Processing Group at the Department of Information Systems of Eotvos Lorand University. Research fields of the group include neurobiological foundation of AI and AI applications.

Gabor Szirtes has been doing his Ph.D. studies in informatics in Andras Lorincz’s group at Eotvos Lorand University, Budapest. He took good place at National Competition in Chemistry and Biology for High School Students. He graduated in chemistry and besides studied physics as well. His major interests are brain modelling, ANN, Independent Component Analysis and its applications.

Balint Takacs finished his M.Sc. studies in biophysics last year and he has been doing his M.Sc. studies in computer science and mathematics in Andras Lorincz’s group at Eotvos Lorand University, Budapest. His major interests are neurobiology, brain modelling and ANN.