BMC Bioinformatics

Poster presentation

Model of visual attention for video sequences

Mariofanna Milanova

Address: Department of Computer Science, University of Arkansas at Little Rock, Little Rock, AR 72204, USA

Email: Mariofanna Milanova - mgmilanova@ualr.edu

This abstract is available from: http://www.biomedcentral.com/1471-2105/9/S7/P14

Background

In our previous work [1] we generalized Olshausen's algorithm [2] and designed a perceptual learning model using video sequences. In this study we propose to model conjunction search. Conjunction search (search for a unique combination of two features – e.g., orientation and spatial frequency – among distractions that share only one of these features) examines how the system combines features into perceptual wholes. We propose to improve the effectiveness of the decomposition algorithm by providing classification awareness. Attentional guidance does not depend solely on local visual features, but must also include the effects of interactions among features. The idea is to group together filters that will be responsible to extract similar features. It is well known that knowledge about which features define the target improves search performance and/or accuracy [3]. The nearest neighbors of the fixations will share a certain feature.

Methods

The main goal of this work is to use sequences of images and to design an attention model including conjunction search based on unsupervised self-learning. First, Independent Component Analysis algorithm is used to determine an initial set of basis functions from the first image (Figure 1). Second, unsupervised self-organizing learning is used to group together similar bases functions (Figure 2).

Conclusion

It is shown that performing sparse learning codes on video sequences of natural scenes produces results with qualitatively similar spatio-temporal properties of simple receptive field of neurons. The basic functions are similar to those obtained by sparse learning, but in our model they have a particular order (Figure 3). The proposed framework was tested using neurobiological (event related potentials ERP's) and behavioral (eye tracking) data.
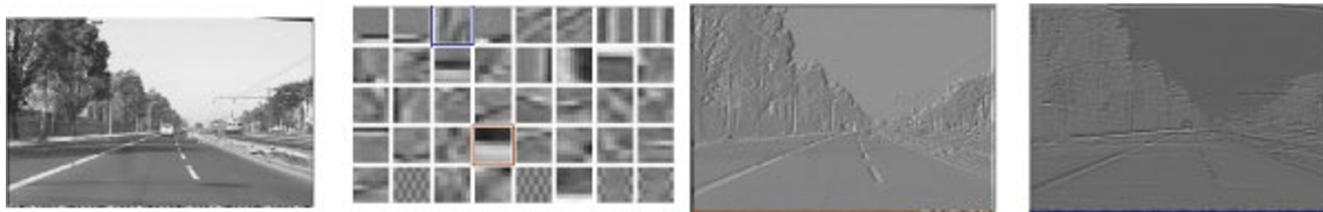

Figure 1
Initial basis images. a) the original image; b) a set of basis functions received by ICA from several patches from one image; c) and d) the convolution results of two functions with the original image.
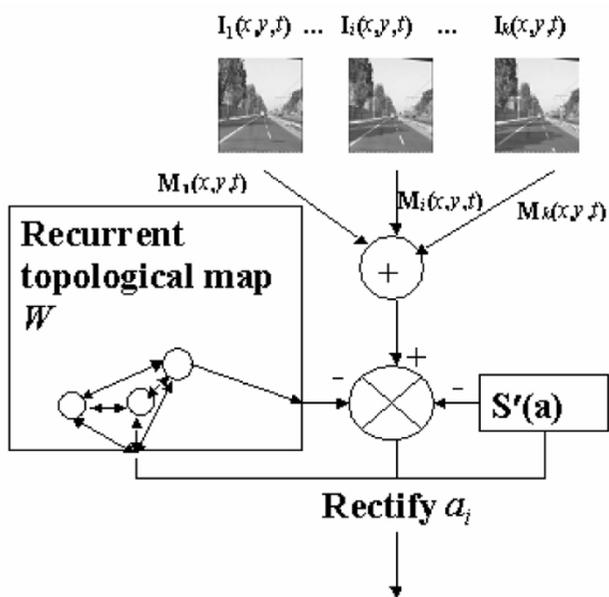
Figure 2
Components for unsupervised self-organizing learning.

References

1. Milanova M, Wachowiak M, Rubin S, Elmaghraby A: **A perceptual learning model based on topological representation, neural networks.** *Proceedings, IJCNN'01 International Conference* 2001:406-411.
2. Olshausen B: **Sparse codes and spikes.** In *Probabilistic Models of Perception and Brain Function* Edited by: *Rao RPN, Olshausen BA, Lewicki MS. MIT Press*; 2001.
3. Rutishauser U, Koch C: **Probabilistic modeling of eye movement data during conjunction search via feature-based attention.** *Journal of Vision* 2007, **7(6):**5. 1–20
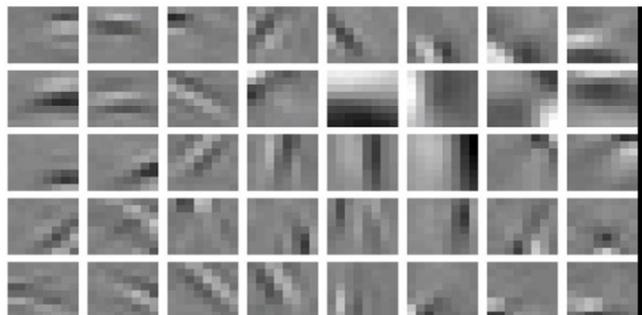
Figure 3
Ordered resulting basis functions.