EEG signals decomposition: a multi-resolution analysis

Decomposizione di segnali EEG: un'analisi multi-risoluzione

Piercesare Secchi, Simone Vantini, and Paolo Zanini

Abstract In this work we analyze a Blind Source Separation (BSS) problem for the EEG signals of patients affected by alcoholism. We solve the BSS problem through an innovative algorithm, named Hierarchical Independent Component Analysis (HICA), able to provide a multi-resolution non-orthogonal data-driven basis. Moreover we show the improvements obtained by HICA in terms of the phenomenological interpretation of the components with respect to other popular BSS techniques.
Abstract *In questo lavoro viene preso in considerazione un tipico problema di Blind Source Separation (BSS) applicato ad un dataset relativo a segnali EEG di pazienti affetti da alcolismo. Il problema è risolto attraverso un algoritmo innovativo, chiamato Hierarchical Independent Component Analysis (HICA), in grado di fornire una base non ortogonale, data-driven e multi-risoluzione. In particolare vengono mostrati i miglioramenti ottenuti dal HICA in termini di interpretazione fenomenologica delle componenti rispetto ad altre tecniche note in letteratura.*

Key words: Blind Source Separation, Independent Component Analysis, multi-resolution, EEG.

Piercesare Secchi
MOX - Department of Mathematics, Politecnico di Milano, piazza Leonardo da Vinci 32, 20133 Milano, Italy e-mail: piercesare.secchi@polimi.it

Simone Vantini
MOX - Department of Mathematics, Politecnico di Milano, piazza Leonardo da Vinci 32, 20133 Milano, Italy e-mail: simone.vantini@polimi.it

Paolo Zanini
MOX - Department of Mathematics, Politecnico di Milano, piazza Leonardo da Vinci 32, 20133 Milano, Italy e-mail: paolo.zanini@polimi.it

1 Introduction

In the analysis of high-dimensional and complex data, the attention is often focused on two connected issues: dimensional reduction (possibly data-driven) and a meaningful synthetic description related to the physics of the problem under study. We look for a reduced space where to represent data. In particular we search for a basis whose elements are representative of the most important features of the phenomenon analyzed. Some of these may involve a great number of the original variables while others may be restricted only to a few. Hence a multi-scale analysis is desirable. Multi-resolution is a crucial property in several applicative scenarios. In particular we consider the analysis of EEG signals of subjects affected by alcoholism. For each patient in the study, measurements from 61 electrodes placed on the scalp are available. The electrodes are located at standard sites [5]. For each electrode, the recorded signal measures the electrode electric potential with respect to some reference electrode and describes the electrical activity of the brain in the neighborhood of the electrode across time. In particular we analyze the brain signals related to one patient.

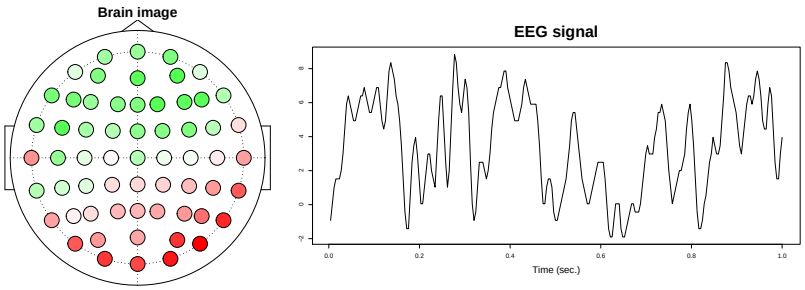

Fig. 1 On the left: EEG brain signals at a fixed instant of time. On the right: EEG for a specific electrode along time.

The subject was exposed to two stimuli. Specifically, the patient was shown two pictures chosen from the 1980 Snodgrass and Vanderwart set [8]. The two stimuli were presented in a matched condition (i.e., the subject has been asked to look at the same picture twice). For each electrode, we observe the signal at 256 equally spaced instants along a time span of 1 second (see Figure 1). We assume that time indexes the sample of observations of a random vector whose components are indexed by electrodes. Hence we consider a sample of $n = 256$ realizations of a random vector $\mathbf{X}$ in $\mathbb{R}^p$, whith $p = 61$. Moreover for the component x_{ij} of $\mathbf{x}_i$, which represents the observed signal of the electrode j at time instant i, we assume the following model:

$$x_{ij} = s_{i1}a_{j1} + \ldots + s_{iK}a_{jK},$$

where $s_{i1},...,s_{iK}$ are the components of a latent source vector $\mathbf{s}_i \in \mathbb{R}^K$, while $a_{j1},...,a_{jK}$ represent unknown real coefficients. This model fits in the well known Blind Source Separation (BSS) framework. If the rows of the $n \times p$ matrix $\mathbb{X}$ collect n observed realizations $\mathbf{x}_1,...,\mathbf{x}_n \in \mathbb{R}^p$ of the random vector $\mathbf{X}$, while the rows of the $n \times K$ matrix $\mathbb{S}$ represent the corresponding unobserved realizations of the latent random source vector $\mathbf{S}$, a BSS problem can be written as

$$\mathbb{X} = \mathbb{S}A^T, \tag{1}$$

where A is a $p \times K$ matrix of unknown coefficients. A BSS problem consists in estimating A and $\mathbb{S}$, given $\mathbb{X}$. In the EEG analysis, the columns of A contain the spatial maps of the brain associated to some reference elements generating the whole signal. Multi-resolution is a compelling property for these maps. Indeed, some brain processes could involve the whole brain, while others activities involve only a specific part of the brain. Imposing multi-resolution property on the elements of A implies the potential for the identification of a wide range of different behaviors characterizing the brain activity. In this work we solve the BSS through Hierarchical Independent Component Analysis (HICA), an innovative method for the construction of a multi-scale non-orthogonal data-driven basis (see [7] for a detailed description of the algorithm).

In section 2 we describe the main ideas driving the HICA algorithm, while in section 3 we present the EEG case study in order to highlight the improvements due to this new approach in terms of phenomenological interpretation of the components.

2 Methods

We here briefly describe a new approach based on the integration between two BSS techniques: Independent Component Analysis (ICA) [3] and Treelets [4]. While ICA assumes stochastic independence between the sources, Treelets provides an estimate of the matrix A through a multi-resolution orthonormal data-driven basis. Even if in many circumstances orthogonality is a desirable property, it often introduces an artificial constraint not related to the phenomenological characteristics of the analyzed problem. The HICA algorithm is a hierarchical procedure that at each level l of the tree provides a multi-resolution non-orthogonal data-driven basis. Specifically, after estimates for $\mathbb{S}$ and A are initialized, the algorithm iterates for $p-1$ times the following three steps:

1. find the two most similar variables according to a suitable measure of dependence (i.e., the distance correlation [9]);
2. compute an Independent Component Analysis of the variables found in step 1 and replace these variables with the new variables identified by the ICA transformation. In this step data, as usual in ICA, are standardized before performing the ICA algorithm;

3. order the two new variables according to their variances. The variable with the lowest variance is stored and, at the next step, only the variable with the greatest variance is considered as a possible candidate for a new aggregation.

At each iteration l we obtain a new estimate for the basis matrix A, say $\widehat{A}^{(l)}$, and a new estimate for the source matrix $\mathbb{S}$, say $\widehat{\mathbb{S}}^{(l)}$. Strategies for choosing the optimal estimates $\widehat{A}^{(l)}$ and $\widehat{\mathbb{S}}^{(l)}$ as well as selecting the appropriate dimension K for the source space, are discussed in [7], where we also prove some consistency properties of the HICA algorithm.

3 EEG analysis

In this section we analyze through HICA EEG signals for a single subject affected by alcoholism. The dataset is courtesy of the online UCI Machine Learning Repository [1]. We show how the properties characterizing the HICA solution, allow to obtain noticeable improvements in terms of phenomenological interpretation. All the analyses are carried out with the fastHICA R package [6] we developed to implement HICA. We compare the results obtained by HICA with those provided by Treelets, ICA (exploiting the fastICA algorithm [3]), and PCA. In Figure 2, for each method we show the five most important basis elements (i.e., $K = 5$) related to the analyzed subject. In details, for PCA we show the first 5 principal components, for ICA the results obtained with the fastICA algorithm selecting 5 sources, while for Treelets and HICA we select the level $l = 55$ and show the 5 components found by the criterium described in [7]. By inspecting Figure 2, it is clear how multi-resolution methods provide localized basis elements. These highlight components defined on localized brain regions and allow to identify more precisely the areas involved in the task. PCA and ICA, instead, provide more unspecific components, possibly difficult to read. This is apparent in the fourth row of Figure 2, where HICA and Treelets select a single electrode (i.e., a single variable). This electrode clearly represents some noise either related to facial muscles activity or due to an unexpected saturation of the electrode. Another interesting component that shows the benefits due to a multi-resolution approach can be seen in the first row of Figure 2. This component is captured both by HICA and Treelets and identifies the associative activity in the frontal brain area which is devoted to elaborate the information related to similarities and differences between images. This crucial component is not caught by PCA and ICA. The second and the fifth row in Figure 2, instead, point out the differences between HICA and Treelets. While Treelets yield an unclear result, with components involving the entire occipital cerebral hemisphere, HICA splits this information into two separate components. The HICA component shown in the second row is related to the primary visual cortex, the first area reached by visual information, which analyzes it in terms of shape and pattern recognition. Then the information flow goes to the internal area of the occipital hemisphere, which associates to the stimulus specific features like color, direction or origin. This area is

highlighted in the fifth row. This separation of the occipital area is caught only by HICA.

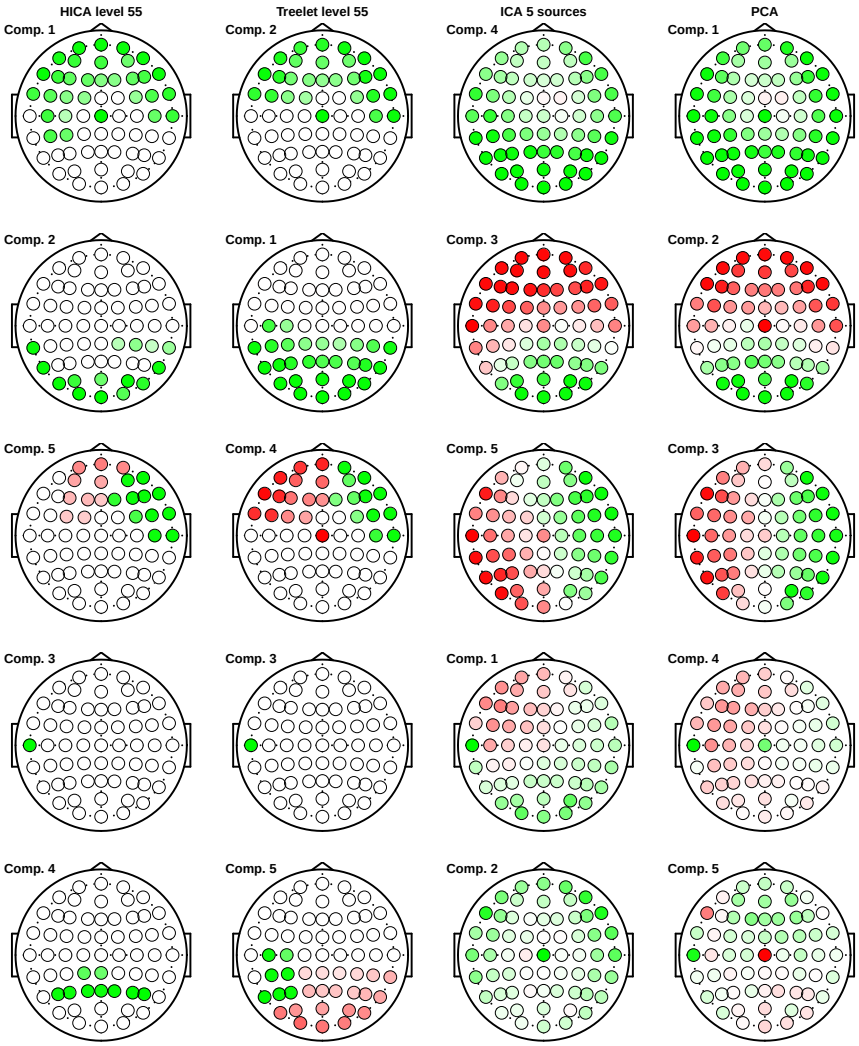

Fig. 2 First five loadings found by HICA (first column on the left), Treelets (second column), ICA (third column), and PCA (fourth column).

References

1. Bache, K., Lichman, M.: UCI Machine Learning Repository. Irvine, CA: University of California, School of Information and Computer Science. http://archive.ics.uci.edu/ml (2013)

2. Comon, P., Jutten, C.: Handbook of Blind Source Separation: Independent Component Analysis and Applications. Academic Press (2010)
3. Hyvarinen, A., Oja, E.: Independent Component Analysis: Algorithms and Applications. Neural Networks, **13**, 411-430 (2000)
4. Lee, A. B., Nadler, B., Wasserman L.: Treelets - an adaptive multi-scale basis for sparse unordered data. The Annals of Applied Statistics, **2**, 435-471 (2008)
5. Sharbrough, F., Chatrian, G.E., Lesser, R.P., Lders, H., Nuwer, M., Picton, T.W.: American Electroencephalographic Society Guidelines for Standard Electrode Position Nomenclature. Journal of Clinical Neurophysiology, **8**, 200-202 (1991)
6. Secchi, P., Vantini, S., Zanini, P.: fastHICA - Hierarchical Independent Component Analysis: a multi-scale sparse non-orthogonal data-driven basis, r package version 1.0 (2014a)
7. Secchi, P., Vantini, S., Zanini, P.: Hierarchical Independent Component Analysis: a multi-resolution non-orthogonal data-driven basis. MOX report 01/2014, MOX - Department of Mathematics, Politecnico di Milano. (2014b)
8. Snodgrass, J.G., Vanderwart, M.: A standardized set of 260 pictures: norms for the naming agreement, familiarity, and visual complexity. Journal of Experimental Psychology: Human Learning and Memory, **6**, 174–215 (1980)
9. Szkely, G.J., Rizzo, M.L., Bakirov, N.K.: Measuring and testing dependence by correlation of distances. The Annals of Statistics, **35**, 2769-2794 (2007)