NeuroImage 42 (2008) 87–98

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Understanding brain connectivity from EEG data by identifying systems composed of interacting sources

Laura Marzetti [a,b,*], Cosimo Del Gratta [a,b], Guido Nolte [c]

[a] Department of Clinical Sciences and Bioimaging, Gabriele D'Annunzio University, Italy
[b] Institute for Advanced Biomedical Technologies, Gabriele D'Annunzio University Foundation, Italy
[c] Fraunhofer FIRST.IDA, Berlin, Germany

ARTICLE INFO

Article history:
Received 14 January 2008
Revised 18 April 2008
Accepted 24 April 2008
Available online 6 May 2008

Keywords:
Electroencephalography
Magnetoencephalography
Principal component analysis
Independent component analysis
Source interaction
Inverse methods

ABSTRACT

In understanding and modeling brain functioning by EEG/MEG, it is not only important to be able to identify active areas but also to understand interference among different areas. The EEG/MEG signals result from the superimposition of underlying brain source activities volume conducted through the head. The effects of volume conduction produce spurious interactions in the measured signals. It is fundamental to separate true source interactions from noise and to unmix the contribution of different systems composed by interacting sources in order to understand interference mechanisms.

As a prerequisite, we consider the problem of unmixing the contribution of uncorrelated sources to a measured field. This problem is equivalent to the problem of unmixing the contribution of different uncorrelated compound systems composed by interacting sources. To this end, we develop a principal component analysis-based method, namely, the source principal component analysis (sPCA), which exploits the underlying assumption of orthogonality for sources, estimated from linear inverse methods, for the extraction of essential features in signal space.

We then consider the problem of demixing the contribution of correlated sources that comprise each of the compound systems identified by using sPCA. While the sPCA orthogonality assumption is sufficient to separate uncorrelated systems, it cannot separate the individual components within each system. To address that problem, we introduce the Minimum Overlap Component Analysis (MOCA), employing a pure spatial criterion to unmix pairs of correlates (or coherent) sources. The proposed methods are tested in simulations and applied to EEG data from human μ and α rhythms.

Introduction

Non-invasive high temporal resolution functional imaging methods, such as electroencephalography (EEG) and magnetoencephalography (MEG), are well suited to study brain dynamics. Their millisecond temporal resolution can be exploited, in fact, not only to follow the variation of the activation patterns in the brain but also to track interference phenomena among brain areas. Recently, much attention has been paid to this second aspect (David et al., 2004; Horwitz, 2003; Lee et al., 2003), and many techniques aimed at studying brain connectivity have been developed and applied based on EEG potentials (Brovelli et al., 2002; Gevins, 1989) and MEG fields (Taniguchi et al., 2000; Gross et al., 2001). Frequency-domain methods are particularly attractive in this sense since the activity of neural population is often best expressed in this domain (Astolfi et al., 2007; Gross et al., 2001; Pfurtscheller and Lopes da Silva, 1999). In the present work, a frequency domain approach is developed in which the imaginary part of the data cross-spectra at a frequency of interest is used to isolate true interference phenomena from the measured data. The imaginary part of the data cross-spectra is, in fact, the only part of complex cross-spectra that reflects true non-zero-lagged interactions in the brain. Hence, this quantity is insensitive to zero-lagged volume conduction effects (Nolte et al., 2004; Marzetti et al., 2007) that can be particularly problematic when describing interdependencies between signals (Gross et al., 2001; Nunez et al., 1997; Nunez et al., 1999). It can be shown that non-interacting sources do not contribute systematically to the imaginary parts of cross-spectra; therefore, this parameter is well suited to study brain interference phenomena (Nolte et al., 2004).

On the other hand, EEG and MEG measure a signal resulting from the superposition of the underlying brain source activities. In order to disentangle the contribution of the various brain

* Corresponding author. Institute for Advanced Biomedical Technologies (ITAB), Gabriele D'Annunzio University, Via dei Vestini, 31, 66013 Chieti, Italy. Fax: +39 08713556930.

E-mail addresses: laura.marzetti@gmail.com, lmarzetti@unich.it (L. Marzetti).

1053-8119/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2008.04.250

sources to the measured signal, methods for the extraction of essential features based on linear transformations of the data space have been widely used. Among them, the well-known second-order methods such as principal component analysis (PCA) (Jolliffe, 1986) and factor analysis (FA) (Kendall, 1975) as well as methods based on higher order statistics such as independent component analysis (ICA) (Hyvarinen et al., 2001). We aim here at identifying the contribution of various sources and/or systems to the interaction phenomenon by decomposing the imaginary part of the cross-spectra. To this end, strategies for the identification of systems composed by interacting sources are developed in order to improve the understanding of the interaction phenomenon. We apply the proposed methods to simulated data and to real data from human μ and α rhythms.

Materials and methods

Problem formulation

We will first introduce some terminology used in this paper. By "source" we refer to functionally diverse neuronal units that are active in a specific rest or task condition. Each source has an activation time course and a specific spatial pattern in a sensor array (i.e., a specific orientation in the "sensor space," in which each axis corresponds to the signal in one sensor). In this paper, we do not consider explicit time dependence as is possible, e.g., in event-related experiments. To study non-stationary effects, the necessary adjustments and re-interpretations are in principle straight forward but the details are beyond the scope of this paper.

It is assumed that the data were measured in a task-related experiment, i.e., processes are not explicitly time dependent. It is possible to generalize the concepts also to event-related experiments, but the necessary adjustments and re-interpretations are beyond the scope of this paper.

Let $x(t)$ be the M-dimensional EEG/MEG data vector at time t. In this paper, we assume that these data can be decomposed as

$$\mathbf{x}(t) = \sum_{p=1}^{P} \left[\alpha_p(t)\mathrm{a}_p + \beta_p(t)\mathrm{b}_p\right] + \sum_{q=1}^{Q} \gamma_q(t)\mathrm{c}_q \tag{1}$$

where $\alpha_p(t)$, $\beta_p(t)$ and $\gamma_q(t)$ are the temporal activities of each source and a_p, b_p, and c_q are the respective spatial patterns. $\gamma_q(t)\mathrm{c}_q c_q$ are Q independent components, and $\alpha_p(t)\mathrm{a}_p$ and $\beta_p(t)\mathrm{b}_p$ are P pairwise interacting components, i.e., $\alpha_p(t)$ can be statistically dependent on $\beta_p(t)$ but neither on $\alpha_r(t)$ nor on $\beta_r(t)$ for $p \neq r$, and analogously for $\beta_p(t)$. We additionally assume that $P \leq M/2$ but we do not make restrictions on Q, the number of independent components. As a "compound system," "interacting system," or simply "system," we refer to the activities denoted by a specific index p in the first sum in Eq. (1). A compound system hence consists of two temporal activities $\alpha_p(t)$ and $\beta_p(t)$, which are associated with two spatial patterns in channel space, a_p and b_p, respectively.

This model makes two assumptions. First, all interactions are assumed to be pairwise. For more complex interactions, the methods outlined below will approximate these systems by dominating pairs of interaction. Second, the number of compound systems measurable with EEG/MEG is assumed to be lower or equal than half the number of channels. Note that this assumption is considerably weaker than the typical assumption used for ICA, namely, that the total number of sources (which include, e.g., channel noise) is not larger than the number of channels. Both assumptions are made for technical reasons. A generalization to arbitrary structure is welcome but is beyond the scope of this paper.

The goal of this paper is to identify the spatial patterns and to localize the sources of the interacting systems by analyzing cross-spectral matrices estimated from the data. Conceptually, this goal is achieved in three steps: (a) we analyze only imaginary parts of cross-spectra to get rid of systematic contributions from non-interacting sources; (b) we introduce a new method (sPCA) and recall an older one (PISA) both capable of identifying the 2D subspaces spanned by a_p and b_p for all p; and (c) we introduce a new method (MOCA) to identify the patterns a_p and b_p themselves for a given 2D subspace. MOCA also includes the localization of the sources.

Principal component analysis and source principal component analysis (sPCA)

PCA is an orthogonal linear transformation that finds a representation of the data using only the information contained in the covariance matrix of the (zero mean) data vector. A new coordinate system for the data is chosen by PCA so that the largest variance by any projection of the data lies on the first axis (first principal component), the second largest variance on the second axis, and so on.

Again, let $x=(x_1,\ldots, x_m)^{\mathrm{T}}$ be an M-dimensional EEG/MEG instantaneous data vector with mean value defined by $\mu_x = E\{\mathbf{x}\}$. The notation $(\cdot)^{\mathrm{T}}$ indicates matrix transpose.

Setting $x_0 = x - \mu_x$, the zero-lag covariance matrix for that data set is defined as

$$\mathbf{C}_x = E\left\{\mathbf{x}_0\mathbf{x}_0^{\mathrm{T}}\right\} \tag{2}$$

Then, finding the principal components is equivalent to finding the eigenvalues and the eigenvectors of the covariance matrix by solving the eigenvalue equation:

$$\mathbf{C}_x\mathbf{U} = \mathbf{U}\boldsymbol{\Gamma} \tag{3}$$

where $\boldsymbol{\Gamma}$ is a diagonal matrix of eigenvalues $\boldsymbol{\Gamma} = \mathrm{diag}\,\{\lambda_1,\ldots,\lambda_m\}$, and the columns of $\mathbf{U}$ are the corresponding eigenvectors that form an orthogonal basis of the data space. For the calculation in the Fourier domain, the analogue of the covariance matrix is the cross-spectra matrix that is defined as

$$\mathbf{C}_x(f) = E\left\{\hat{\mathbf{x}}_0(f)\hat{\mathbf{x}}_0^{\mathrm{H}}(f)\right\} \tag{4}$$

where $\hat{x}_0(f)$ is the Fourier transform of the Hanning-windowed demeaned data for each trial, $E\{\cdot\}$ is estimated by the average over these trials, and $(\cdot)^{\mathrm{H}}$ denotes the Hermitian conjugate, i.e., transpose and complex conjugate.

When interpreting the eigenvectors of such matrices as potentials of single sources, two assumptions are made: (a) the sources are uncorrelated, and (b) the potentials are mutually orthogonal in signal space. Let's assume that the first assumption is correct and address the second one. The forward mapping of sources to sensors is essentially a spatial low pass filter. Even if the regions of non-vanishing source activities are spatially distinct, such as dipoles at different locations, which are orthogonal as vector fields, the respective potentials are highly blurred and are in general spatially correlated. Apparently, the orthogonality assumption would be much more reasonable if we were able to reconstruct the sources and formulate this assumption in source space.