

Independent component analysis of fMRI group studies by self-organizing clustering

Fabrizio Esposito,^{a,b,c,*} Tommaso Scarabino,^b Aapo Hyvarinen,^d Johan Himberg,^e
Elia Formisano,^f Silvia Comani,^g Gioacchino Tedeschi,^a Rainer Goebel,^f
Erich Seifritz,^h and Francesco Di Salle^{b,c}

^aSecond Division of Neurology, Second University of Naples, Italy

^bDepartment of Neuroradiology, Scientific Institute “Casa Sollievo della Sofferenza”, S. Giovanni Rotondo, Italy

^cDepartment of Neurological Sciences, University of Naples “Federico II”, Italy

^dHIIT Basic Research Unit, Department of Computer Science, University of Helsinki, Finland

^eNeural Networks Research Centre, Helsinki University of Technology, Finland

^fDepartment of Cognitive Neuroscience, Maastricht University, The Netherlands

^gDepartment of Clinical Sciences and Bio-imaging, University of Chieti, Italy

^hUniversity Hospital of Clinical Psychiatry, University of Bern, Switzerland

Received 14 July 2004; revised 21 October 2004; accepted 28 October 2004
Available online 8 January 2005

Independent component analysis (ICA) is a valuable technique for the multivariate data-driven analysis of functional magnetic resonance imaging (fMRI) data sets. Applications of ICA have been developed mainly for single subject studies, although different solutions for group studies have been proposed. These approaches combine data sets from multiple subjects into a single aggregate data set before ICA estimation and, thus, require some additional assumptions about the separability across subjects of group independent components. Here, we exploit the application of similarity measures and a related visual tool to study the natural self-organizing clustering of many independent components from multiple individual data sets in the subject space. Our proposed framework flexibly enables multiple criteria for the generation of group independent components and their random-effects evaluation. We present real visual activation fMRI data from two experiments, with different spatiotemporal structures, and demonstrate the validity of this framework for a blind extraction and selection of meaningful activity and functional connectivity group patterns. Our approach is either alternative or complementary to the group ICA of aggregated data sets in that it exploits commonalities across multiple subject-specific patterns, while addressing as much as possible of the intersubject variability of the measured responses. This property is

particularly of interest for a blind group and subgroup pattern extraction and selection.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Functional magnetic resonance imaging; fMRI; Exploratory data-driven analysis; Descriptive statistics; Multidimensional projection; Spatial; Temporal similarity analysis; Independent component analysis; Group analysis; Multi-subjects analysis

Introduction

Independent component analysis (ICA) is a valuable tool for the multivariate data-driven analysis of functional magnetic resonance imaging (fMRI) data (McKeown et al., 1998a,b, 2003). As a purely data-driven methodology, ICA does not require the specification of temporal signal profiles or anatomical regions of interest to generate meaningful spatiotemporal patterns of brain activity. The multivariate statistical nature of ICA allows one to transform three-dimensional fMRI data sets into brain activity patterns starting from the spatial or temporal covariance of the measured signals and reveals multiple spatiotemporal “modes” of signal variability (Friston et al., 1993). This transformation is achieved by imposing the general, yet neurophysiologically plausible, constraint of removing the statistical dependence of the output modes (Brown et al., 2001; McKeown et al., 1998a). In order to meet this constraint, the value distribution of the fMRI signals in space or time is to be considered: the variant called spatial ICA (sICA) refers to the statistical distribution of signals across the sampled hemodynamic locations, while the variant called temporal ICA

* Corresponding author. Department of Neurological Sciences, University of Naples Federico II, II Policlinico (Nuovo Policlinico) Padiglione 17, Via S. Pansini 5, 80131 Naples, Italy. Fax: +39 081 546 3663.

E-mail address: faesposi@unina.it (F. Esposito).

Available online on ScienceDirect (www.sciencedirect.com).

(tICA) refers to the statistical distribution of source signals across the sampled time-points (Calhoun et al., 2001a).

Both sICA and tICA have been used in different contexts. The tICA is applied to fMRI measurements in the same way, and with the same assumptions that ICA is commonly applied to EEG or MEG recordings; on the other hand, the structure of whole-brain three-dimensional fMRI data sets has suggested the sICA as the default ICA variant for fMRI.

The neurological significance of applying sICA for the decomposition of single-subject fMRI time-series can be seen in the two equivalent formulations of ICA. First, the modes of signal change separated by ICA are such that the mutual information is minimized, that is, each generated pattern carries a minimum amount of information about the other patterns. Second, the statistical distribution of the sources are maximally far from the Gaussian distribution (Hyvarinen et al., 2001). In fact, the first definition extends the concept of functional connectivity patterns of brain imaging data (Friston et al., 1993), where multiple brain regions are unified by their time-courses, with the constraint that none of these regions systematically occurs in two different patterns. The second definition fits with the concept of “activation map”, for which the amount of functional information is related to how the values of a few (active) voxels are significantly different from the remaining (Gaussian-distributed, as default) “mass” of voxels: the more Gaussian the distribution of a three-dimensional map is, the less selective and, thus, functionally uninformative, the resulting pattern will be.

Based on this theoretical background, ICA has been successfully used for the decomposition of individual fMRI time-series. However, since fMRI studies increasingly involve the statistical comparison of more than one group of subjects, for example, healthy people vs. people with a disease, it has become necessary to develop strategies to extending the ICA analysis framework from single-subject to group studies and multi-group studies.

The most natural and intuitive way that avoids additional assumptions for the individual ICA data model is to perform fixed- or random-effects analyses on the results of the decompositions of each individual data sets (Calhoun et al., 2001b; Seifritz et al., 2002). The main challenge of this approach is to integrate the ICA analysis chain with a suitable post-estimation analysis step in which an automatic tool would allow a systematic matching of the estimated components across all the subjects of the study. However, in previous studies applying ICA to fMRI, the matching of the component maps was based on subjective and context-specific criteria: in the absence of general and effective tools for the subject- or group-level selection of “matching” components, this approach remains difficult to implement, and the loss of sensitivity caused by a possible mismatch of components cannot be easily corrected.

Conventional model-driven univariate methods (e.g., regression analysis) have been naturally generalized from single- to multi-subject methods by simple schemes of across-subject data aggregation based on matrix averaging or concatenation. Previous work has proposed similar schemes to combine the individual data sets into a single group data set prior to performing one single ICA run on a group data-matrix. Two alternative approaches have been proposed. Following the typical matrix notation, they can be referred as column-wise (or subject-wise) (Calhoun et al., 2001c), and row-wise (across time-courses) concatenation (Svendsen et al., 2002). These methods have been reviewed and compared using artificial data sets to the simplest

across-subject averaging by Schmithorst and Holland (2004). In order to be correctly applied, both approaches require the substantial assumption that a given source of signal change exists as an “observable process” in all of the subjects entering the analysis. Specifically, column-wise aggregation imposes a common space of observations for all the sources (the normalized anatomical space), although it allows different activation time-courses for the different subjects. Row-wise aggregation imposes a common time-course for a generic source to all of the subjects, although it allows “no activity” to occur in some of subjects. Despite the additional, sometimes restrictive hypotheses required by the aggregate approaches, the use of a common space of observation may serve as useful “regularization” for the estimation of group components.

After ICA parameter estimation, the separation of subject-specific components is achieved by a subject-level unmixing of group components in the column-wise approach and by a vector disaggregation of group components in the row-wise approach. In a more recent work (Calhoun et al., 2004), a new variant of the column-wise group ICA approach was presented, where single-subject component time-courses were obtained using a spatial multiple regression of the group component images onto the individual fMRI data for each time point.

A further approach is the simple across-subject averaging (Schmithorst et al., 2004): although the computational load is the least extensive and is independent from the number of subjects, it allows group inferences only indirectly through a subsequent conventional general linear model analysis with the estimated ICA mixing matrix acting as a pseudo design matrix in a way similar to that described by McKeown (2000). In all of the three approaches, at least one form of “non-selective” pooling of different subjects’ data is necessary before estimating group components: spatial for column-wise, temporal for row-wise and spatiotemporal in across-subject averaging.

Although the validity of these approaches in producing “single-group” ICA patterns compatible with individual ICA patterns has been demonstrated, it is noteworthy that they cannot easily predict how much bias or loss of sensitivity may occur in the ICA estimation (and, thus, subsequent random effects analysis of the patterns) in the presence of factors affecting the homogeneity across subjects of the components. Thus, the homogeneity of the sample of subjects, which crucially affects the performances of the random effects analysis for model-driven parametric estimates (Friston et al., 1999), appears to be an even more crucial problem in the context of ICA.

The problem of the homogenous presence of sources in different subjects may occur for many different reasons: for instance, Burbaud et al. (2000) show different activation patterns for mental calculation relating to different strategies (verbal or visual), while in the study of Castelo-Branco et al. (2002), an individual could or could not produce a measurable response related to his/her subjective perception of an ambiguous stimulus.

So far, preliminary attempts to examine the homogeneity (or stationarity) of subject (or timepoint-to-timepoint) homogeneity have been presented by Liao et al. (2004) and Calhoun et al. (2001d).

In general, both predictable (e.g., gender, age, etc.) as well as not easily predictable factors may occur, which can bias the group ICA model estimation, but a comprehensive evaluation of this bias and the possible loss of accuracy of the proposed ICA method is not straightforward.

Download English Version:

<https://daneshyari.com/en/article/9198488>

Download Persian Version:

<https://daneshyari.com/article/9198488>

[Daneshyari.com](https://daneshyari.com)