Independent Component Analysis of spatially distributed patterns of brain activation measured by fMRI

S. TAKERKART[1], M. BENHARROSH[1], I. C. DAUBECHIES[1, 2]

[1]CSBMB, Princeton University, Green Hall, Princeton, NJ 08540, USA

[2]PACM, Princeton University, Fine Hall, Princeton, NJ 08540, USA

[takerkar, mbenharr, icd]@princeton.edu

Résumé – Ces dernières années, l'Analyse en Composantes Indépendantes (ICA) a été utilisée avec succès pour détecter l'activité cérébrale lors d'expériences d'Imagerie par Résonance Magnétique fonctionnelle (IRMf). Depuis, la précision des résultats obtenus a été étudiée en utilisant comme critères la localisation des zones d'activités, ou leur évolution temporelle, à la fois sur des données simulées et réelles. D'autres travaux récents ont démontré l'importance de la distribution spatiale de l'intensité de l'activité à l'intérieur d'une région d'intérêt de grande taille. Nous présentons dans cet article des résultats prouvant que ICA peut estimer avec précision ces « figures d'activité distribuée », en examinant des données IRMf simulées et réelles.

Abstract *– In the last few years, Independent Component Analysis (ICA) has proven to be useful in detecting brain activation in functional Magnetic Resonance Imaging (fMRI) experiments. Since then, its accuracy has been studied both in terms of the location of the detected activation kernels, and of the temporal behaviors, on both simulated datasets and real experiments with simple paradigms. Recently, the relevance of a new kind of information, the spatial distribution of the magnitude of activation in large « distributed patterns », has been demonstrated. In this paper, we first present a new way to create simulated fMRI datasets containing such distributed patterns of activation ; we then show that ICA algorithms can accurately estimate these patterns, both on simulated datasets and on real data from the original study.*

1. Introduction

Functional Magnetic Resonance Imaging (fMRI) is a tool that measures changes in the blood oxygenation level across time. Traditional data analysis methods are strongly model-driven [1]. They are suitable for numerous experimental designs in which the paradigm is thought to give a strong basis for hypothesis testing. But these methods lack the capability to detect the unpredictable (« what if something happens in the brain that does not follow the temporal evolution dictated by the task ? »). This has led researchers to investigate the use of data-driven analysis methods, such as Independent Component Analysis (ICA). Since its initial application [2], the circumstances in which ICA succeeds at detecting and estimating fMRI activation, and the robustness of these results, have been studied by various groups [3], [4]. In this paper, we study the performances of ICA algorithms to estimate « distributed patterns of activation », which were first described in [5]. After detailing the model used to apply ICA on fMRI datasets, we describe the characteristics of the datasets used in this study, and then present the results, comparing them with the standard model-based approach.

2. ICA model for fMRI data analysis

Four-dimensional (3D in space + time) fMRI datasets can be represented as a 2D matrix **X**, of size **TxV**, where **T** is the number of time-points and **V** the number of voxels included in the analysis. Hence, each row **Xi** contains the intensity values of all voxels of the brain at time point **i**. When applied in the spatial domain, ICA decomposes the **Xi** as a linear combination of **N** statistically independent spatial maps. The problem is therefore to estimate a linear model **X=MC,** where **C** is an **NxV** matrix (each of its rows representing a spatial map) and **M** is a **TxN "**mixing" matrix (each of its columns containing the time-course of its associated map). We used two algorithms, widely described in the literature, namely Infomax [6], which minimizes the mutual information between spatial maps, and FastICA [7], which maximizes the negentropy. Principal Components Analysis (PCA) is used as an initialization step (whitening) and to reduce the dimensionality of the problem from **T** to **N,** by keeping only the first **N** principal components.

3. Data

3.1 Simulation model

We introduce a new way to simulate fMRI datasets, so that the simulated datasets contain « distributed patterns of activation ». We first acquired a null dataset: a set of EPIs (echo-planar images) while the subject had no task to perform. We then added some activation to this dataset in several steps. We first chose the anatomical region onto which activation would be super-imposed: this was done by acquiring an auxiliary functional time-series during which the same subject was presented with a flashing checkerboard in an on/off paradigm. This set of images was co-registered with the null dataset and the visual cortex was identified by using

a standard General Linear Model (GLM); it represented approximately one tenth of the brain voxels present in the volume. This ensured that the simulated activation would be added in an anatomically correct region (gray matter). Secondly, a unique boxcar time-course was added to all voxels contained in this region. The original aspect of this method was that the magnitude of activation randomly varied from voxel to voxel, between 0.5% and 3.5% of signal change from the baseline level, thus creating a "distributed pattern of activation". One hundred simulated datasets were generated with different randomly generated patterns, for each of which the true pattern of activation was recorded to allow for evaluation later on.

3.2 Real experiment

We then used real data from an experiment described in [5]: the authors measured with fMRI the brain activation from six subjects viewing pictures of faces, cats, five categories of man-made objects (bottles, shoes, scissors, chairs, houses) and control non-sense images (scrambled). Twelve time-series were recorded for each subject. Each time-series was designed with eight 24 second stimulus blocks (one for each category), separated by 12 second intervals of rest. The object-selective cortex was defined by the voxels whose responses differed significantly by category. Patterns of response were examined within the ventral temporal object selective cortex. In order to examine the consistency of the spatial patterns, the data was split into two sets whether the time-series were odd or even numbered (thus constituting a training dataset and a test dataset). In the original analysis, a GLM defined by sixteen predictors (each of them modeling the response to a given category relative to rest, for each set, odd or even) was used to determine the pattern of response to each category: the sixteen resulting _ maps weighing each predictor were used as an estimate of the strength of the category-specific patterns of response.

4. Data analysis

All datasets were motion corrected using AIR [8]. When data chunks from several time-series were concatenated (as explained later), a linear detrending and an adjustment of the mean of each time series was conducted, to avoid baseline jumps. Then ICA was applied on intracranial voxels, selected with a brain extraction algorithm described in [9], after a whitening and dimensionality reduction step executed with PCA.

The selection of the component of interest for a given experimental condition was done in several steps: 1. building a reference function (which is on for the trials of the condition of interest, off otherwise); 2. ranking the components according to the correlation coefficient of their time-course with this reference function; 3. selecting the highest one. Although this selection process did require information about the experimental paradigm, the estimation of the model was fully data-driven. This post-hoc sorting of the components was carefully validated before being applied on all datasets, by verifying that it agreed on numerous and various datasets with a visual assessment of all component maps and their associated time-course.

The accuracy of the estimation of distributed patterns of activation was evaluated by computing Pearson correlation coefficients between pairs of patterns : a pattern known a priori (the true recorded pattern for the simulations, or the pattern estimated on the training dataset for the real data), and a pattern estimated from the test dataset

When processing real data, we compared patterns estimated in two experimental conditions : this produced two estimates (from the test and training datasets) for each of the two categories. Category-specific patterns estimated from the test dataset were compared to the ones estimated on the training dataset in the following way: for each pair of categories (A, B), four correlations scores were computed between the four available patterns: $A_{training}$, A_{test}, $B_{training}$ and B_{test}. If the within-category correlation was larger than the between-category correlation the comparison was counted as correct identification, thus proving the consistency of the spatial distribution of the magnitude of activation for a given condition.

Two ICA algorithms were used : Infomax, as implemented in the code available at http://www.cnl.salk.edu/~enghoff/, and FastICA, as implemented in the FSL package (Melodic : http://www.fmrib.ox.ac.uk/fsl/melodic2/).

All the results were compared with the ones produced when using a standard GLM, where the category-specific patterns were estimated with the weight coefficients (_ maps) of the corresponding predictor in the regression.

5. Results and discussion

5.1 Simulations

The averaged correlation scores across all one hundred simulated datasets are shown in Table 1. The results were good overall : all scores were greater than 0.79. The very low standard errors show the robustness of the estimation to the shape of the pattern. Among the two algorithms tested, Infomax performed better than FastICA in all cases. In most cases, ICA algorithms outperformed the GLM. We believe this is due to the fact that ICA does not rely on any assumption as to the shape of the hemodynamic response, and that it allows a better implicit modeling of noise sources : indeed, among the **N** components estimated, only one contains « activation », whereas the **N-1** others contain information about other sources, which can be related to the motion of the subject's head (even after motion correction), or to sources of instrumental noise (EPI ghosting, scanner drifts etc.)

The influence of dimensionality reduction was studied by modifying the number of independent components **N** to be estimated, which was done classically by keeping the **N** first principal components in a pre-processing step. The choice of the number of components presents a problem for the researcher who uses ICA to analyze fMRI data since no standard way to make this choice exists. A general rule of

thumb which emerged from the literature is to choose **N** between one quarter and one third of the number of time points **T**. Recently, a maximum-likelihood criterion was proposed to estimate the number of sources actually present in an fMRI dataset [10]. This criterion suggested to use **N = 30** components for all our simulated datasets, which coincides with the empirical rule mentioned above, since we had **T=104**. For both algorithms, the correlation scores at **N=30** are lower than at **N=75** or **N=104** (the difference being statistically significant at p=0.07 for Infomax, and p=0.00006 for FastICA, when a one-sided paired t-test was performed). This suggests that for this particular purpose, estimating spatially distributed patterns of activation and evaluating this estimation with a spatial correlation analysis, a too strong dimensionality reduction performed with PCA can harm the estimation.

TABLE 1 : Mean spatial correlation scores (± standard error), between patterns estimated with 1. ICA (for different amounts of dimensionality reduction) and 2. the GLM.

Dimension. reduction	FastICA	Infomax
10 components	0.7971 ± 7.10^{-4}	0.8465 ± 5.10^{-4}
30 components	0.8153 ± 6.10^{-4}	0.8654 ± 5.10^{-4}
50 components	0.8177 ± 6.10^{-4}	0.8662 ± 5.10^{-4}
75 components	0.8187 ± 6.10^{-4}	0.8664 ± 5.10^{-4}
None (104 components)	0.8495 ± 5.10^{-4}	0.8665 ± 5.10^{-4}
GLM	0.8036 ± 7.10^{-4}	

While it is intuitive that performing dimensionality reduction with PCA might throw away some meaningful data (since fMRI activation signals are usually very small compared to the noise level, they are not expected to always be represented in the first principal components), this practice has become standard in the fMRI field, for different types of reasons: practical ones, like result sorting (there is still no standard way to select "meaningful" components); and theoretical ones, since applying PCA results in whitened data, which eases the convergence of ICA algorithms [11]. But this experiment shows that using PCA for dimensionality reduction might actually be harmful, suggesting that other alternatives should be studied.

5.2 Real data

We first applied ICA separately on each original functional time-series, which contains one block of trials for each of the eight categories. None of the components showed a time course with a high correlation score with any of the category-specific reference functions. But one unique consistently task-related component was detected: its time course showed activation period for the task, without any distinction between categories. This map thus describes the mean response across categories. ICA seems to lack sensitivity to distinguish responses across categories. When several original time-series were concatenated (regardless of the number of time-series) into a single ICA analysis, the results were similar. This was to be expected since the original analysis conducted in [5] showed that the patterns of activation associated with each category are highly overlapping. These patterns are thus strongly statistically dependent, which explains why they are not separated by ICA when all were present in a single dataset.

We therefore re-organized the data in a manner that would gather trials from one category in a unique dataset. The time-series were rearranged by concatenating the blocks of images corresponding to the same category (24 seconds of task + the following 12s of rest). This was done separately for the training and the test data-sets (odd- and even-numbered time-series), thus creating sixteen new composite time-series (two for each of the eight categories), each containing six blocks of stimulus of a unique category. For each of these, ICA retrieved one CTR component. We were then able to compute the pairwise correlation scores within- and between-category (shown on Figure 1), and the identification accuracy rate (shown on Table 2). In both cases, we compared our results obtained with ICA to the ones originally published in [5], where the distributed patterns were estimated with a GLM analysis.

The results presented in Figure 1 show that, when estimated with ICA, within-category correlation scores between patterns are in most cases higher than between-category scores. This is confirmed by the computation of category identification presented in Table 2, which are above chance (50 %) in all but one case. This first reinforces the fact that the information contained in the spatial distribution of the magnitude of activation within a region of interest is relevant to brain mapping. This shows too that ICA algorithms are suitable to estimate these patterns, and that they are robust to complex data manipulation (as the re-organization of the functional time-series). As with the simulated data, the Infomax algorithm outperfomed FastICA in all cases, significantly in this case : whereas the results from Infomax compares to the original results with the GLM, FastICA produces lower accuracy.

The influence of dimensionality reduction on the analysis of real data was unclear because of the reorganization of the data. Indeed, when the numbers of chosen components increased, it coincided with the appearance of transiently task-related components which spread the information contained in the distributed patterns into several components. When the number of components was small, the spatial correlation scores globally decreased. The results presented were for a "classically" chosen: **N=30** components.

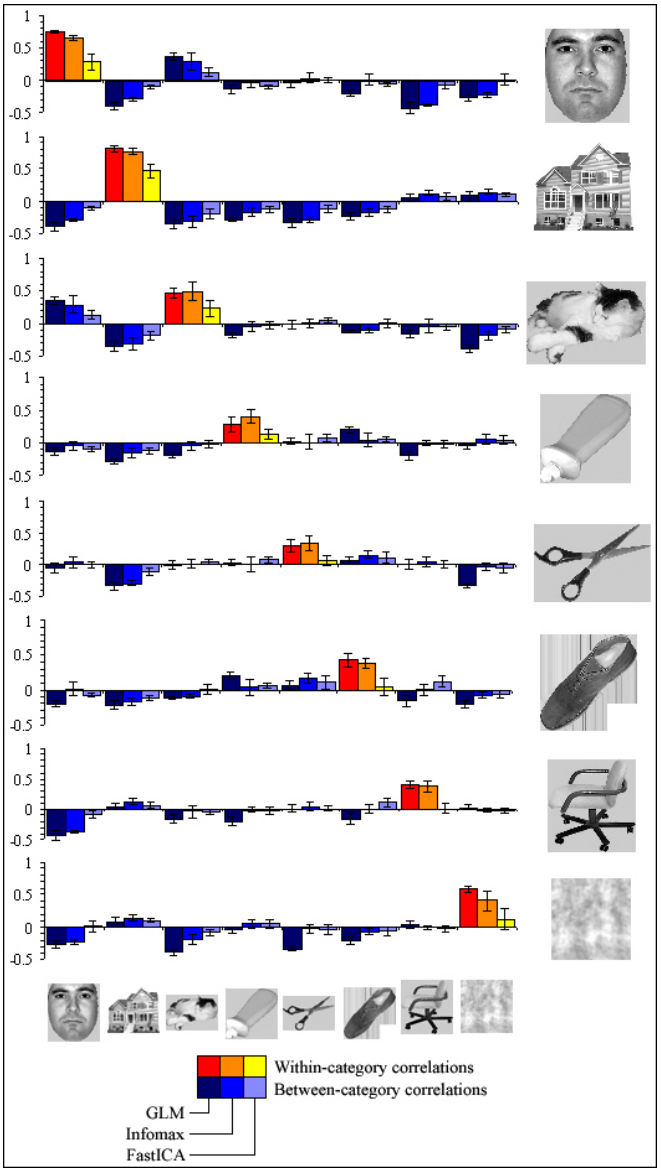

FIGURE 1 : Mean (across subjects) correlation scores (± standard error) for all pairs of categories, in all ventral temporal object selective cortex.

TABLE 2 : Identification accuracy of categories on the test dataset for the GLM (results from [5]) and ICA algorithms.

Category	GLM	FastICA	Infomax
Face	100 %	79 %	100 %
House	100 %	90 %	100 %
Cats	98 %	71 %	88 %
Bottles	90 %	79 %	95 %
Scissors	92 %	57 %	83 %
Shoes	92 %	62 %	95 %
Chairs	96 %	43 %	93 %
Scramble	100 %	64 %	83 %

6. Conclusion

We showed in this study that ICA algorithms are able to estimate accurately spatially distributed patterns of activation measured by fMRI in both simulated and real data and that they outperform traditional model-based analysis methods on our simulated datasets. This confirms the usefuleness of ICA algorithms for fMRI data analysis. However, this study pointed out two weaknesses of these blind source separation methods, which will define directions of future work : the dimensionality reduction method using PCA was shown to be sub-optimal ; and these algorithms cannot separated strongly overlapping patterns of activation.

Acknowledgments

The authors would like to thank James Haxby for sharing his data and for his useful comments.

This work was supported by the NIH grant R01 MH67204.

References

[1] K.J. Worsley and K.J. Friston, *Analysis of fMRI time-series revisited – again.* NeuroImage vol. 2, pp. 173-181, 1995.

[2] M.J. McKeown et al., *Analysis of fMRI data by blind separation into independent spatial components*, Human Brain Mapping vol. 6, pp. 160-188, 1998.

[3] F. Esposito et al., *Spatial component analysis of fMRI time-series : to what extent do results depend on the algorithm used ?* Human Brain Mapping vol. 16, pp. 146-157, 2002

[4] V. Calhoun et al., *Spatial and temporal independent component analysis of fMRI data containing a pair of task-related waveform*, Human Brain Mapping vol. 13, pp. 43-53, 2001.

[5] J.V. Haxby et al, *Distributed and overlapping representations of faces and objects in ventral temporal cortex*. Science vol. 293, pp. 2425-2430, 2001.

[6] A.J. Bell and T.J. Sejnowski, *An information-maximisation approach to blind separation and blind deconvolution*. Neural Computation, vol. 7, pp. 1003-1034, 1995.

[7] A. Hyvarinen, *Fast and robust Fixed-Point algorithms for independent component analysis*. IEEE Transactions on Neural Networks, vol. 10, pp. 626-634, 1997.

[8] R. Woods et al., *Automated Image Registration: I. General Methods and Intrasubject, Intramodality Validation*. Journal of Computer Assisted Tomography, vol. 22, pp.139-152, 1998.

[9] S. Smith, *Fast robust automated brain extraction*, Human Brain Mapping, vol. 17, pp.143-155, 2002.

[10] C. Beckmann, et al., *Investigating the intrinsic dimensionality of FMRI data for ICA*. Proc. Of Seventh Int. Conf. on Human Brain Mapping, 2001.

[11] J.F. Cardoso, *On the performance of orthogonal source separation algorithms*, Proc. EUSIPCO 1994, pp. 776-779