DYNAMICAL COMPONENTS ANALYSIS OF FMRI DATA

Bertrand Thirion, Olivier Faugeras

INRIA Sophia-Antipolis,
Projet Robotvis,
2004 route des lucioles, BP 93, FR-06902 Sophia Antipolis

ABSTRACT

In this article, we present a new multivariate analysis method for the analysis of fMRI data. This method tries to capture the deterministic structure present in the time series, using either an autoregressive scheme or the knowledge of the experimental paradigm, so that the interpretation of the spatio-temporal patterns is achieved in parallel with their detection. In the spatial domain, the components are made maximally independent through an ICA-like criterion. A global criterion is derived to express the model priors as well as the goodness of fit. The method is a priori adaptable to every sort of experimental conditions (block or event-related design). An experiment is presented on real data to show the potential of the method for the detection of signals, the analysis of their content as well as their localization.

1. INTRODUCTION

Multivariate analysis of fMRI data has been considered as an exploratory or descriptive tool only, while proper inference should rather be done in a univariate framework. One straightforward reason is the difficulty of bringing priors into a multivariate framework: Singular value decomposition (SVD or PCA) [1] builds a linear and decorrelated representation of the data from its covariance matrix; clustering methods [2] sum up the data by using similarity of the time courses or derived features; independent component analysis (ICA) [3] builds a linear model with spatio-temporal components that are maximally independent either in the spatial or in the temporal [4] domain. All these methods try to summarize the structure of the data, and can be viewed as preprocessing before another algorithm makes inference of the data [5].

These methods can be considered as bias free, since they impose very weak constraints on the data structure, but they finally bring out little useful information about the data, since they do not take into account the most important information, that is the experimental paradigm. Recently, a multivariate spectral method has been proposed [6], where the information of interest is introduced as the stimulation frequency, but this is useful for periodic paradigms only.

In this paper, we propose a new multivariate method based on the distinction between the random/stochastic signals, and the deterministic/predictable ones. More precisely, we try to extract spatio-temporal components that are maximally predictable in the temporal domain while being spatially mutually independent. The predictability of a time series may appear as a vague concept, but as in [7], one can use a generalization of ARMA processes: the signal can be predicted given its past values (AR model) or given the values of an *excitatory* input (MA model). In a multivariate framework, it seems appropriate to make a choice between the two alternatives, since we want to *unmix* different effects present in the data.

At this step, we can propose a new interpretation of a fMRI dataset: it appears as a dynamical system excited by an experimental paradigm, whose law of evolution has to be explored. This is what our algorithm intends to do.

The paper is organized as follows: in section 2 we describe our model and how we derive an algorithm to process fMRI data. Then we present experiments on real data on section 3, and discuss briefly the results.

2. DYNAMICAL COMPONENTS ANALYSIS

Let $X(t), t = 1..T$ be a set of temporal data ; each $X(t)$ is of size N and can be thought of as an image. A DCA is a decomposition

$$X(t) = \sum_{k=1}^{K} I_k s_k(t) + N(t)$$

where $1 \leq K \leq min(T, N)$; I_k are spatial components of size N. $\forall k,\ \ s_k(t)$ behaves as a dynamical system i.e. it has a deterministic behavior and is indexed by a 1-D (time) variable.

$N(t)$ is purely random noise (in the temporal domain), i.e. it contains all the non deterministic part of the original data.

Work partially supported by INRIA ARC fMRI and European Grant Mapawamo, No QLG3-CT-2000-30161.

Temporal model: Each $s_k(t)$ should behave as *dynamical systems* ; this means that their values is predictable either given the past values $(s_k(t-l)_{l=1..L})$ or given the values of the paradigm $(p(t-m)_{l=0..M})$. The most general equation is

$$s_k(t) = F(s_k(t-1), .., s_k(t-L), p(t), .., p(t-M))$$

Here, we consider the following special case $s_k(t) = F(x_k(t))$ where $x_k(t)$ is a *predictor* of s_k which can be either $\sum_{l=1}^{L} \alpha_l s_k(t-l)$ or $\sum_{m=0}^{M} \beta_m p(t-m)$ or a combination of the two terms.

$x_k(t) = \sum_{l=1}^{L} \alpha_l s_k(t-l)$ means that the signal is autoregressive. This is well suited for periodic oscillators.

$x_k(t) = \sum_{m=0}^{M} \beta_m p(t-m)$ means that the dynamics of the system essentially reduces to a filter of the paradigm variable. This filter can be nonlinear, if F is not an affine function.

We propose to derive the best predictor x_k by maximizing $\prod_{t=1}^{T} P(s_k(t)|x_k(t))$; indeed, we can assume that the deviation from the prediction is not autocorrelated. Taking the log of this expression, this amounts to minimizing $H(s_k|x_k)$, the conditional entropy of s_k for a given x_k. This is equivalent to maximizing the mutual information between x_k and s_k. This is made easier by starting with a linear initialization of x_k.
The choice for L and M is a compromise between goodness of fit and model length. Here we chose a priori $L = M = 5$.

Spatial model The requirement for spatial components is that they should be maximally independent. This amounts to considering the set of 3-D images I_k as realization of an unknown random K-dimensional probability function whose marginals are independent. In a maximum Likelihood framework, its identification has been shown [8] to yield the minimization of the mutual information of the marginals ; adding a decorrelation constraint reduces the criterion to the sum of the marginal entropies of the images $H(I_k)$. Thus, the log-likelihood of our spatial criterion is the well-known ICA criterion :

$$\mathcal{L} = -\sum_{k=1}^{K} H(I_k) + cste$$

Spatio-temporal criterion Here we derive a criterion to obtain I_k and s_k. We follow a ML approach. Let $P(I, s|X)$ be the probability of the model given the data X:

$$P(I, s|X) \sim P(X|I, s).P(I, s)$$

The term in the right-hand side is the probability of observing the data given the model while $P(I, s)$ is the probability of observing the model. As a first step, we choose

$$P(X|I, s) = \frac{1}{(\sqrt{2\pi\sigma^2})^T} \prod_{t=1}^{T} \exp\left(-\frac{||X - \sum_{k=1}^{K} I_k s_k||^2(t)}{\sigma^2}\right)$$

which amounts to assuming that $N(t)$ is gaussian with mean 0 and covariance $\frac{\sigma^2}{N} I_N$, and
$P(I, s) = P(I)P(s)$, which means that the spatio-temporal model factorizes in time and space. We have shown in the previous paragraph that our spatial prior yields
$log(P(I)) = -\sum_{k=1}^{K} H(I_k) + cste$. Next it seems plausible to consider the s_k as independent; thus
$P(s) = \prod_{k=1}^{K} P(s_k)$ so that
$\log P(s) = \sum_{k=1}^{K} \log(P(s_k)) = -\sum_{k=1}^{K} H(s_k|x_k) + cste$, where x_k is the optimal predictor defined above. The right-hand side term is our prior of predictability on each signal. Now,

$$\log(P(I, s|X)) = c - \sum_{t=1}^{T} \frac{||X(t) - \sum_{k=1}^{K} I_k s_k(t)||^2}{\sigma^2} +$$

$$\sum_{k=1}^{K} \frac{1}{T} \sum_{t=1}^{T} \log(P(s_k(t)|x_k(t))) - \sum_{k=1}^{K} H(I_k)$$

Finally, maximizing the likelihood of the model amounts to minimizing the functional

$$J_\sigma(K) = \sum_{t=1}^{T} \frac{||X(t) - \sum_{k=1}^{K} I_k s_k(t)||^2}{\sigma^2} +$$

$$\sum_{k=1}^{K} H(s_k|x_k) + \sum_{k=1}^{K} H(I_k)$$

subject to the constraint that the I_k are orthogonal (i.e. decorrelated).

Practical implementation The algorithm is initialized with the selection of the K first PCA components of the data, which minimizes the residual variance term and enforces the spatial orthogonality constraint. Then the system is optimized by sucessive relaxation of the three energy criteria, and stops when a fixed point is reached. The spatial orthogonality constraint is enforced at each iteration.

Let us notice that it is hopeless to obtain K meaningful time courses; but the quantity that we have defined to quantify the predictability of the component, i.e. $MI(s_k, x_k)$ is a straightforward tool to select in fine the components of interest.

3. EXPERIMENTS AND RESULTS

The data : We present an experiment on real data, which has already been published [9].

The present analysis is reduced to one subject performing the following experiment (called fMRI2 in [9]): A visual stimulation is performed, with 4 conditions: Heading, dimming static, dimming flow, and baseline. The *Heading* condition means that the subject views a ground plane optic flow pattern that simulates self-motion ; *dimming static* is a control task where no self-motion is simulated, but a part of the stimulus display is slightly dimmed, and *dimming flow* is another control condition specially designed to disentangle spatial and featural attention through the dimming of the ground plane flow; during the *baseline* condition the subject fixates without stimulation. The subject had to press a key when he identified a Heading or control task.

The paradigm is blocked with each block lasting 10 scans. The order of conditions is pseudo-randomized. The length of the time series is 720 scans. Each condition is realized during 180 scans.

Results : The only preprocessing performed on the data was a motion correction implemented in SPM, and centering of the data. We then applied our algorithm, which was parametered to produce $K = 80$ dynamic components. To deal with the multiplicity of conditions, we simply used each stimulus time course as regressors for the optimal predictor. The optimal combination of these regressors, gives the relative amount of activation attributable to each condition (see figure 2). After convergence one component was consistently paradigm-related, and the other ones contained either oscillating signals (autoregressive predictor) or no signal of interest (i.e. the mutual information with the predictor was low). In figure 1, we display horizontal views of our spatial components that contain the regions of interest. SPM F-maps of the *effects of interest* obtained with SPM were very similar (data not shown).

The areas of activations are early visual areas (1a-d), hMT/V5+(1c-d), DIPSM/L(1e) and a dorsal premotor area(1f).

The observation of the task-related time course together with the corresponding predictor (figure 2a) gives much insight on the relative activation strength between conditions. This is simply given by the decomposition of the optimal predictor of the signal on the set of the time courses of the different conditions. This relative strength is displayed in part b of figure 2. Once again, these results confirm those obtained in [9]: Heading condition dominates the others, and dimming flow yields activations weaker than heading condition but stronger than dimming static condition.

Discussion : It seems too early to draw meaningful neurophysiological conclusions from our method : Unlike current methods, we do not build here a hypothesis testing framework to assess our findings. This is a matter for future work. The main advantage that is brought so far is that some usual pre-processing steps of fMRI data analysis (e.g. detrending, smoothing, high pass filtering), are not necessary with our method; in spite of its intrinsically exploratory nature, it can extract the relevant information of the dataset. As such, it can thus be a tool to check whether sophisticated methods do not yield artefactual responses. Its adaptability to different experimental designs enables also to give some insight on the content of a dataset before hypothesis testing.

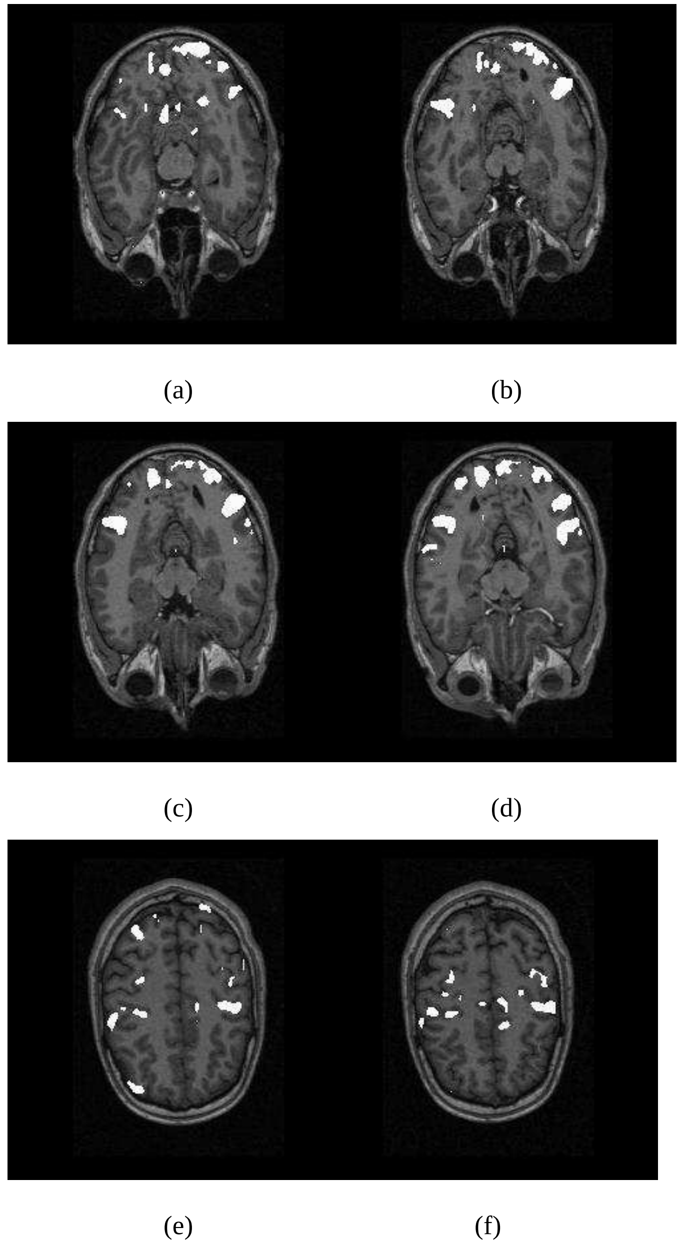
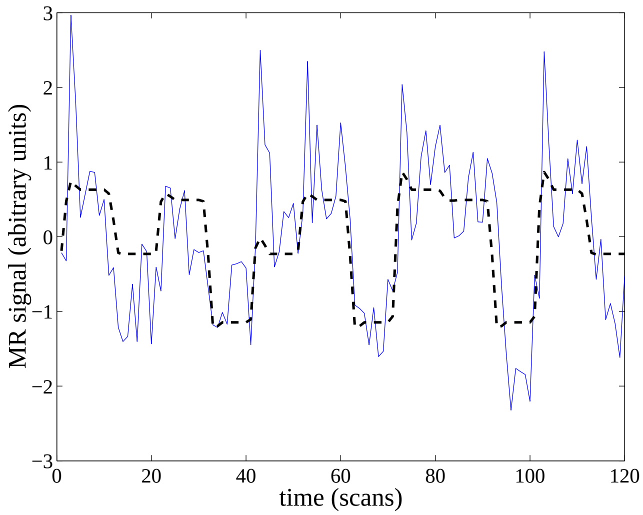
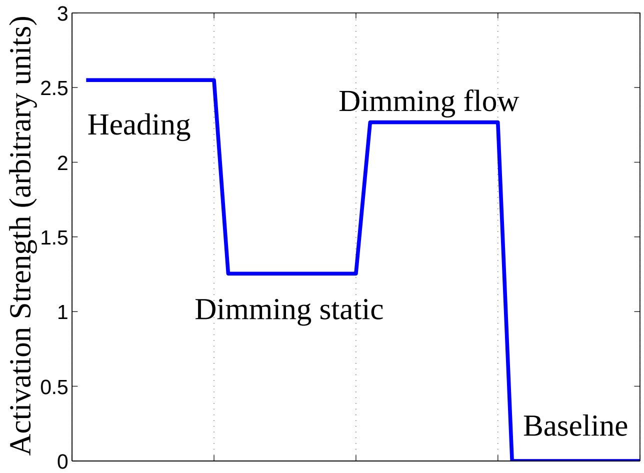


Fig. 1. Spatial activation map for the consistently task-related component. The horizontal slices (a) to (h) span regions where maximal activations were found, from bottom to top of the cortex.

Another question is the advantage with respect to existing multivariate methods, especially ICA. Besides the theoretical advantage of orienting more properly the search of components of interest, we have noticed that our additional criterion and the usual ICA criterion help each other in the optimization process, which means that the conjunction of temporal and spatial priors is useful.



(a)



(b)

Fig. 2. Consistently task-related time course : a) First 120 scans of the time course (continuous line) together with the computed optimal predictor (dotted line). The corresponding sequence of stimuli is (Head = Heading, Dsta = Dimming Static, Dflo = Dimming Flow, Base = Baseline): Head Dsta Dflo Base Dsta Dflo Base Head Dflo Base Head Dsta. b) Relative activation strength of the conditions in the same component. Heading condition yields the strongest activation followed by dimming flow and dimming static.

4. CONCLUSION

Dynamical Component Analysis is a new multivariate analysis of fMRI data, which can bring soft constraints into data analysis. We believe that this is a well-founded way to get more information than usual decomposition or demixing methods. In particular, it seems appropriate to define the components of interest as those that are maximally predictable; these components are precisely those that break usual analysis hypotheses (stationarity, independence, normality). We have shown that the method brings out relevant information on the data, without the use of the usual preprocessing methods.

Besides the necessity of a more systematic tuning of the algorithm parameters, there remains to build a statistical framework to give a more precise assessment of the obtained results.

Acknowledgements

We wish to thank Professor G. Orban, S.Sunaert and H. Peuskens, who provided us with the functional MR images we used. The work was developed in collaboration with the laboratory of Neurophysiology, K.U.Leuven, Medical School, Leuven, Belgium (LEUNEURO), directed by G. Orban.

5. REFERENCES

[1] K.J. Friston, C. D. Frith, R. S. J. Frackowiak, and R. Turner, “Characterizing dynamic brain responses with fmri: a multivariate approach.,” *NeuroImage*, vol. 2, pp. 166–172, 1995.

[2] R. Baumgartner, C. Windischberger, and E. Moser, “Quantification in functional magnetic resonance imaging : fuzzy clustering vs correlation analysis.,” *magnetic resonance Imaging*, vol. 16, no. 2, pp. 115–125, 1998.

[3] Martin J. McKeown, S. Makeig, et al., “Analysis of fmri data by blind separation into independant spatial components,” *Human Brain Mapping*, vol. 6, pp. 160–188, 1998.

[4] V.D. Calhoun, T. Adali, G.D. Pearlson, and J.J. Pekar, “Spatial and temporal independent component analysis of functional mri data containing a pair of task-related waveforms,” *Human Brain Mapping*, vol. 13, pp. 43–53, 2001.

[5] Martin J. McKeown, “Detection of consistently task-related activations in fmri data with hybrid independent component analysis,” *NeuroImage*, vol. 11, pp. 24–35, 2000.

[6] Karsten Müller, Gabriele Lohmann, Volker Bosh, and D. Yves von Cramon, “On multivariate spectral analysis of fmri times series,” *NeuroImage*, vol. 14, pp. 347–356, 2001.

[7] J.W. Fisher, E.R. Cosman, C. Wible, and W.M. Wells, “Adaptative entropy rates for fmri time-series analysis,” in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2001*. 2001, vol. 2208 of *Lecture Notes in Computer Science*, pp. 905–912, Springer.

[8] Jean-Francois Cardoso, “Blind signal separation: statistical principles,” in *Proceedings of the IEEE. Special issue on blind identification and estimation*, Oct. 1998, vol. 9, pp. 2009–2025.

[9] H. Peuskens, S. Sunaert, P. Dupont, P. Van Hecke, and G.A. Orban, “Human brain regions involved in heading estimation,” *The Journal of Neuroscience*, vol. 21, no. 7, pp. 2451–2461, Apr. 2001.