Helsinki University of Technology
Publications in Computer and Information Science

Report E7
December 2006

FUNCTIONAL ELEMENTS AND NETWORKS IN fMRI

Jarkko Ylipaavalniemi Eerika Savia Ricardo Vigário Samuel Kaski

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY
TECHNISCHE UNIVERSITÄT HELSINKI
UNIVERSITE DE TECHNOLOGIE D'HELSINKI

Distribution:
Helsinki University of Technology
Department of Computer Science and Engineering
Laboratory of Computer and Information Science
P.O. Box 5400
FI-02015 TKK, Finland
Tel. +358-9-451 3267
Fax +358-9-451 3277

This report is downloadable at
http://www.cis.hut.fi/Publications/

ISBN-13 978-951-22-8570-9
ISBN-10 951-22-8570-3
ISSN 1796-2803

Functional Elements and Networks in fMRI

Jarkko Ylipaavalniemi[1], Eerika Savia[1,2], Ricardo Vigário[1] and Samuel Kaski[1,2] *

1- Helsinki University of Technology - Adaptive Informatics Research Centre
2- Helsinki University of Technology - Helsinki Institute for Information Technology
P.O. Box 5400, FI-02015 TKK - Finland

Abstract. We propose a two-step approach for the analysis of functional magnetic resonance images, in the context of natural stimuli. In the first step, elements of functional brain activity emerge, based on spatial independence assumptions. The second step exploits temporal covariation between the elements and given features of the natural stimuli to identify functional networks. The networks can have complex activation patterns related to common task goals.

1 Introduction

Functional magnetic resonance imaging (fMRI) is one of the most successful methods for studying the living human brain. Traditionally, its analysis relies on artificially generated stimuli, coupled with generic statistical signal processing, in clear hypothesis-driven setups. The rising interest in natural stimuli studies calls for the development of new processing approaches.

Completely data-driven methods, when applied to natural uncontrolled stimuli, will discover all active brain processes, regardless of the study's research questions. In this work we propose a two-step approach, where independent component analysis (ICA) finds spatially independent functional elements, whereas nonparametric dependent component analysis (DeCA[1]) collects them into networks related to the natural stimulation. The processing framework is schematically shown in Figure 1.

Related canonical correlation approaches have been previously suggested for fMRI [1, 2, 3]. Yet, no method has dealt with natural stimuli, since all experiments rely on non-overlapping block designs. Furthermore, the novel functional framework introduced in this paper extends the interpretation ability of the analyses.

2 Material and Methods

The experiments use a dataset from a recent competition organized by the University of Pittsburgh [4]. The data consists of fMRI recordings of three subjects as they viewed short film clips and a set of features describing the stimulus.

*This work was partly supported by the IST Programme of the European Community, under the PASCAL Network of Excellence, IST-2002-506778. This publication only reflects the authors' views.

[1]DeCA generalizes canonical correlation analysis (CCA).

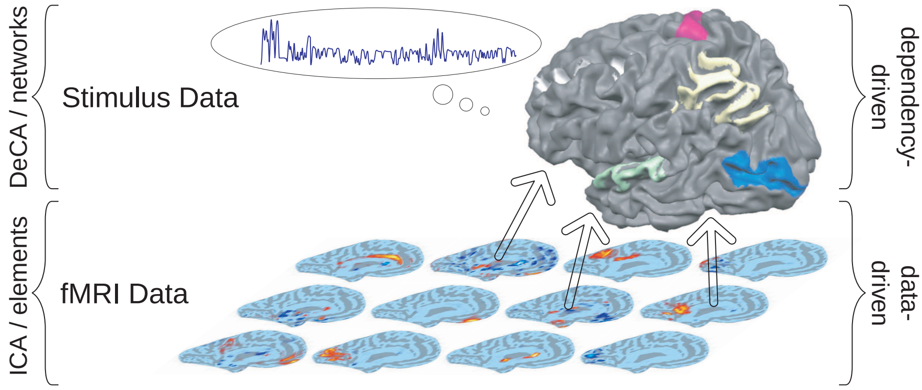


Fig. 1: The proposed framework: elements of functional brain activity emerge from the data via ICA. Functional networks are revealed by DeCA, based on covariation between the elements and task goals, encoded as features.

2.1 Natural Stimulus fMRI Recordings

Data related to the first subject viewing the first movie clip was analyzed. It consists of 20 minutes of continuous fMRI measurements. Whole head volumes were acquired with a 3T scanner using an EPI sequence (TR=1.75s, TE=25ms, slice=3.5mm, FOV=210mm, flip=76°), resulting in 64×64×34 voxels per volume, for 858 time points.

Preprocessing provided by the competition organizers included motion correction, slice time correction, linear trend removal, and spatial normalization of the volume data. Then the cortical surface was extracted and morphed into a smooth inflated surface containing 238735 vertices. Additionally, we retained only 641 time points of the data that contained actual movie viewing.

2.2 Features of Natural Stimulus Data

The film clips, *i.e.*, the natural stimuli, were described with 29 features in the dataset. Some of the features were quantitative, such as *brightness* or *rms sound*, measuring image intensity and root-mean-square sound amplitude, respectively. On the other hand, most of the features were more qualitative, *e.g.*, *laughter* and *sadness*, based on subjective labelings given by the viewers. There were strong correlations among the 29 features. From strongly correlating pairs of features we always selected only one to the experiment. Brain activity related to observing other people's actions is of particular interest, so we combined the original actor-specific features (*e.g.*, *al* and *brad*) into a single new feature *people* by taking their maximum value. We left out features related to places (*e.g.*, *kitchen* and *backyard*).

The resulting set of 9 features was *attention*, *brightness*, *faces*, *food*, *language*, *laughter*, *rms sound*, *sadness* and *people*, shown in Figure 2. Each feature was normalized to have zero mean and unit variance.

Fig. 2: The 9 features of natural stimuli as functions of time.

2.3 Independent Component Analysis

Independent component analysis (ICA) [5] is one of the most popular methods for solving the blind source separation (BSS) problem in a purely data-driven manner. BSS consists of finding solutions to the mixture $\mathbf{X} = \mathbf{AS}$, where only the observed data $\mathbf{X}$ is known. ICA assumes only statistical independence of sources $\mathbf{S}$, and full rank of mixture $\mathbf{A}$. Independence is considered here in the spatial domain, and the mixing reveals the temporal activation patterns of the corresponding sources [6].

Here we use a reliable ICA approach, proposed in [7], based on multiple runs of FastICA [5], in a bagging framework, *i.e.*, with resampled data and randomized initializations.

Suitable parameter values for FastICA were selected heuristically, based on performance, overfitting avoidance, and computation requirements. We used *tanh* nonlinearity in *symmetric* mode looking for 25 independent components in the initially 50 dimensional whitened space. The bootstrapping used a sampling of 25% with correlation threshold of 0.8 and power 4 (see [7] for details on implementation). The reliable ICA included 100 runs and mean representatives of the 25 most reliable components were selected as potential functional elements.

2.4 Nonparametric Dependent Component Analysis

Our goal is to find the underlying factors that are common for two paired datasets, the independent components of brain activity and the features of natural stimuli. Obviously, in both datasets there is also variation that is not shared and we try to distinguish between common and unshared variations. A classical method for this task is the canonical correlation analysis (CCA) [8], which has been shown to maximize mutual information for Gaussian data [9]. In this study the data is far from Gaussian, so we used nonparametric dependent component analysis (DeCA) [10]. It maximizes a dependency estimate between the two datasets y and z:

$$f(\mathbf{w}_y, \mathbf{w}_z) = \sum_{i=1}^{N} \log \frac{q_{yz}(i)}{q_y(i) q_z(i)} \quad ,$$

where $\mathbf{w}$ are the parameters that define the components (linear projection directions) and q denote Parzen estimators of the density of the projected data. The summation is taken over the samples. Intuitively, if the two projected datasets would be near to independent, the estimator of joint density q_{yz} would roughly factorize to the product $q_y q_z$ of the marginals, resulting in zero cost. DeCA looks for projections that are as far from this situation as possible. Multiple components were computed in a deflation manner, looking for one component at a time and removing the found component from the data before looking for the next one.

We run DeCA 20 times from different initializations and took the best solutions according to the cost function. If the next best projections had the same kind of profile we averaged the solution over them.

3 Results

Consistent with previous studies, the reliable independent components in Figure 3 represent spatially independent functional regions of the brain. For example, component with index 3 (IC3) corresponds to the sensory auditory areas in the superior temporal lobes of both hemispheres, whereas IC20 corresponds to the anterior cingulate gyrus. Interpretations for all the components is out of the scope of this paper.

Two illustrative functional networks identified by the method are shown in Figure 4, corresponding to the first dependent component, and Figure 5, corresponding to the fourth dependent component. In both cases, 6 independent components with the highest loading values are shown, sorted according to the absolute loading values. High absolute loading value means that the corresponding IC or stimulus feature contributes significantly to the respective dependent component. The first network comprises areas corresponding to, *e.g.*, auditory (IC3), visual (IC12), and multi-modal integration (IC24). This suggest that the functional role of the network is related to combining information from many sensory inputs. Indeed, the four highest scoring features of the dependent component are *attention*, *people*, *brightness* and *language*. The second network includes areas related to, *e.g.*, language processing (IC3 and IC5) and face recognition (IC8). Additional information can again be drawn from the four highest scoring features *language*, *faces*, *laughter* and *attention*.

4 Discussion

We introduced a two-step approach to the identification of networks of functional brain activity. The method was tested on fMRI recordings of brain responses to natural stimuli. The found networks seem plausible, considering the very subjective and unreliable nature of the available goal settings. Different networks partially shared individual elements, each with a clear functional contribution to the network's common goal. More controlled studies are being planned to verify the results and to further develop the approach.

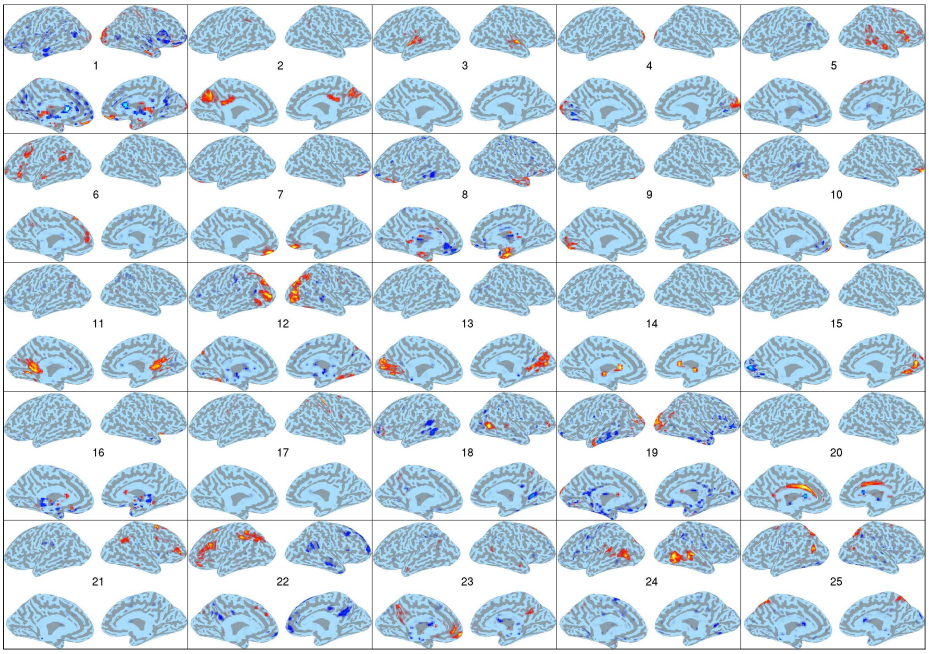
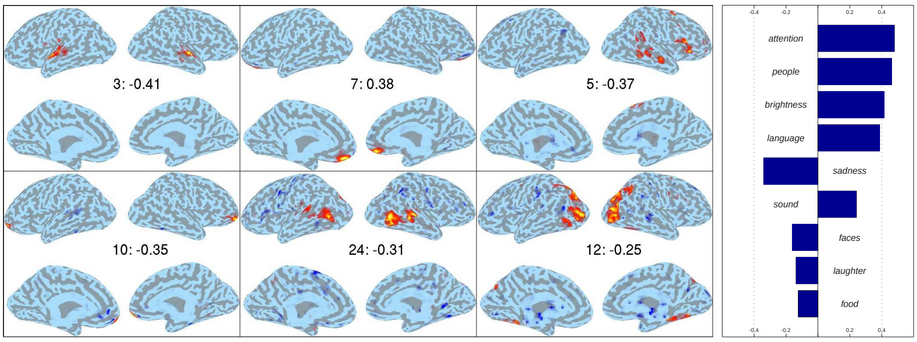
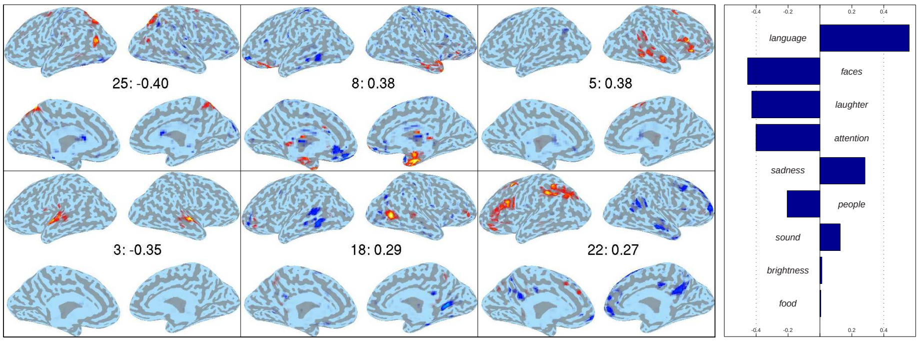


Fig. 3: Overview of 25 independent components representing spatially independent functional brain regions. The index of each component is shown in the middle of each square. Lateral and medial views of both hemispheres of the inflated cortex are shown. The cortex anatomy in light gray shades and the activation pattern superimposed with dark (yellow and blue gradients in color version).



(a) (b)

Fig. 4: (a) The 6 ICs (from Fig. 3) corresponding to the highest loadings in the first dependent component. The loading value is shown in the middle of each square. (b) Respective loadings of the stimulus features.



(a) (b)

Fig. 5: (a) The 6 ICs (from Fig. 3) corresponding to the highest loadings in the fourth dependent component. (b) Respective loadings of the stimulus features.

References

[1] O. Friman et. al. Detection of neural activity in functional MRI using canonical correlation analysis. *Magnetic Resonance in Medicine*, 45(2):323–330, 2001.

[2] T. Youssef et. al. Robust ordering of independent components in functional magnetic resonance imaging time series data using Canonical correlation analysis. In *Proceedings of the SPIE Medical Imaging*, volume 5031, pages 332–340, 2003.

[3] D. R. Hardoon, J. Mour ao Miranda, M. Brammer, and J. Shawe-Taylor. Unsupervised fMRI Analysis. In *NIPS Workshop on New Directions on Decoding Mental States from fMRI Data*, Whistler, Canada, 2006.

[4] W. Schneider et. al. Competition: Inferring Experience Based Cognition from fMRI. In *Proceedings of the 12th Meeting of the OHBM*, Florence, Italy, 2006. http://www.ebc.pitt.edu/2006/competition.html.

[5] A. Hyvärinen et. al. *Independent Component Analysis*. Wiley-Interscience, New York, NY, 2001.

[6] M. J. McKeown et. al. Analysis of fMRI Data by Blind Separation Into Independent Spatial Components. *Human Brain Mapping*, 6(3):160–188, 1998.

[7] J. Ylipaavalniemi and R. Vigário. Analysis of Auditory fMRI Recordings via ICA: A Study on Consistency. In *Proceedings of the IJCNN*, pages 249–254, Budapest, Hungary, 2004.

[8] N. H. Timm. *Applied Multivariate Analysis*. Springer, New York, NY, 2002.

[9] S. Kullback. *Information Theory and Statistics*. Dover Publications, Mineola, NY, 1997.

[10] A. Klami and S. Kaski. Non-parametric Dependent Components. In *Proceedings of the ICASSP*, pages 209–212, Philadelphia, PA, 2005.