ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 509-514

Model–Free Functional MRI Analysis Using Topographic Independent Component Analysis

Anke Meyer–Bäse, Thomas Dan Otto,

Department of Electrical and Computer Engineering, Florida State University,
Tallahassee, Florida 32310-6046, USA

Thomas Martinetz,
Institute for Neuro– and Bioinformatics, Luebeck University,
Luebeck 23569, Germany

Dorothee Auer, and
Max Planck Institute of Psychiatry, Munich 80804, Germany

Axel Wismüller
Department of Clinical Radiology, Ludwig–Maximilians University,
Munich 80336, Germany.

Abstract. Model–free fMRI analysis methods are either transformation–based like principal component analysis (PCA) and independent component analysis (ICA), or based on unsupervised clustering. Since each of these two methods has its particular strengths, it is natural to look for an unifying paradigm of these analysis methods. This is given by the topographic independent component analysis. While achieved by a slight modification of the ICA model, it can be at the same time used to define a topographic order between the components, and thus has the usual computational advantages associated with topographic maps. In this contribution, we can show that when applied to fMRI analysis it outperforms powerful representants of both model–free methods.

1. Introduction

Functional magnetic resonance imaging with high temporal and spatial resolution is a potential method to map rapid and fine activation patterns of the human brain. According to both theoretical estimations and experimental results, an activated signal variation appears very low on a clinical scanner. Thus, analysis methods are required to find the response waveforms and associated activated regions. Generally, these methods can be divided into two categories depending on whether or not they require prior knowledge about activation patterns: model-based and model-free. To eliminate the bias and limitation of model-based analysis methods and to satisfy the demand to analyze data with complicated experimental conditions, analysis methods that do not rely on any assumed model of functional response are necessary. There are two kinds of model-free methods. The first method, principal component analysis (PCA) or independent component analysis (ICA), transforms original data into high-dimensional vector space to separate functional response and various noise sources from each other.

The second method, fuzzy clustering analysis or self-organizing map, attempts to classify time signals of the brain into several patterns according to temporal similarity among these signals.

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 509-514

Since each of these methods has its particular strengths, it is natural to look for an unifying model that combines topographic mapping with ICA. This is given by the topographic independent component analysis [HH01]. Achieved by a slight modification of the ICA model, it can be at the same time used to define a topographic order between the components, and has thus the usual computational advantages associated with topographic maps.

2. The Topographic ICA Architecture

The paradigm of topographic ICA has its roots in [HH00] where a combination of invariant feature subspaces [Koh96] and independent subspaces [Car98] is proposed.

To introduce a topographic representation in the ICA model, it is necessary to relax the assumption of the independence among neighboring components s_i. This makes it necessary to adopt an idea from self–organized neural networks, that of a lattice. It was shown in [HH01] that a representation which models topographic correlation of energies is an adequate approach for introducing dependencies between neighboring components.

In other words, the variances corresponding to neighboring components are positively correlated while the other variances are in the wide sense independent. The architecture of this new approach is shown in Figure 1.

This idea leads to the following representation of the source signals

$$s_i = \sigma_i z_i \tag{1}$$

where z_i is a random variable having the same distribution as s_i and, the variance σ_i is fixed to unity.

The variance σ_i is further modeled by a nonlinearity:

$$\sigma_i = \phi\left(\sum_{k=1}^{n} h(i,k)u_k\right) \tag{2}$$

where u_i are the higher order independent components used to generate the variances, while ϕ describes some nonlinearity. The neighborhood function $h(i,j)$ can be either a two–dimensional grid or have a ring–like structure. Further u_i and z_i are all mutually independent.

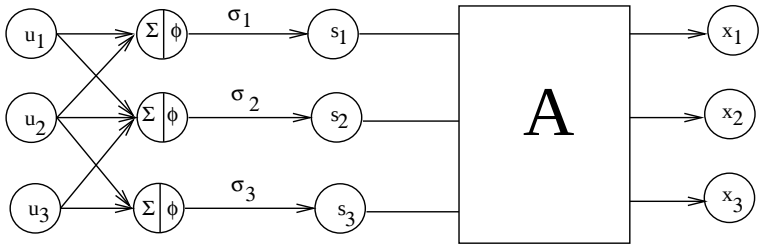


Figure 1: Topographic ICA model [HH01]. The variance generated variables u_i are randomly generated, and mixed linearly inside their topographic neighborhoods. This forms the input to nonlinearity ϕ, thus giving the local variance σ_i. Components s_i are generated with variances σ_i. The observed variables are x_i are obtained as with standard ICA from the linear mixture of the components s_i.

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 509-514

The learning rule is based on the maximization of the likelihood. First, it is assumed that the data is preprocessed by whitening and that the estimates of the components are uncorrelated. The log–likelihood is given by:

$$\log L(\mathbf{w_i}, i = 1, \cdots, n) = \sum_{t=1}^{T} \sum_{j=1}^{n} G(\sum_{i=1}^{n} (\mathbf{w_i}^T \mathbf{x}(t)^2) + T \log |\det \mathbf{W}| \qquad (3)$$

The update rule for the weight vector $\mathbf{w}_i$ is derived from a gradient algorithm based on the log–likelihood assuming $\log |\det \mathbf{W}| = 0$:

$$\Delta \mathbf{w}_i \propto E\{\mathbf{x}(\mathbf{w}_i^T \mathbf{x}) r_i\} \qquad (4)$$

where

$$r_i = \sum_{k=1}^{n} h(i,k) g(\sum_{j=1}^{n} k(k,j)(\mathbf{w}_j^T \mathbf{x})^2) \qquad (5)$$

The function g is the derivative of $G = -\alpha_1 \sqrt{u} + \beta_1$. After every iteration, the vectors $\mathbf{w}_i$ in equation (4) are normalized to unit variance and orthogonalized. This equation represents a modulated learning rule, where the learning term is modulated by the term r_i.

The classic ICA results from the topographic ICA by setting $h(i,j) = \delta_{ij}$.

3. Results and Discussion

FMRI data were recorded from six subjects (3 female, 3 male, age 20-37) performing a visual task. In five subjects, five slices with 100 images (TR/TE=3000 / 60msec) were acquired with five periods of rest and five photic simulation periods with rest. Simulation and rest periods comprised 10 repetitions each, i.e. 30s. In one subject eight slices with 64 images (TR/TE=4000/66msec) were obtained starting with a period of rest lasting 10 repetitions (i.e. 40s) followed by three periods of stimulation alternating with periods of rest comprising 9 repetitions each, i.e. 36 s. Resolution was $3 \times 3 \times 4$ mm. The slices were oriented parallel to the calcarine fissure. Photic stimulation was performed using an 8 Hz alternating checkerboard stimulus with a central fixation point and a dark background with a central fixation point during the control periods [WLD+02]. The first scan was discarded for remaining saturation effects. Motion artifacts were compensated by automatic image alignment (AIR, [WCM92]).

The obtained results were evaluated by (1) assessment of cluster assignment maps, (2) task–related activation maps and (3) associated time–courses. Cluster assignment maps represent cluster membership maps obtained based on a minimal distance criterion in the pixel time course space. For the fMRI data, a comparative quantitative evaluation among the topographic ICA and a traditional ICA, the Infomax, and a clustering method, "neural gas" network, is going to be performed.

3.1. Topographical ICA Results

The clustering results for the topographical ICA are shown in Figure 2. Figure 2 shows the independent time courses emerging from topographical ICA. We immediatley can see a topographical representation by looking at the fifth row from top: five in amplitude similar artifactual time courses are grouped together.

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 509-514

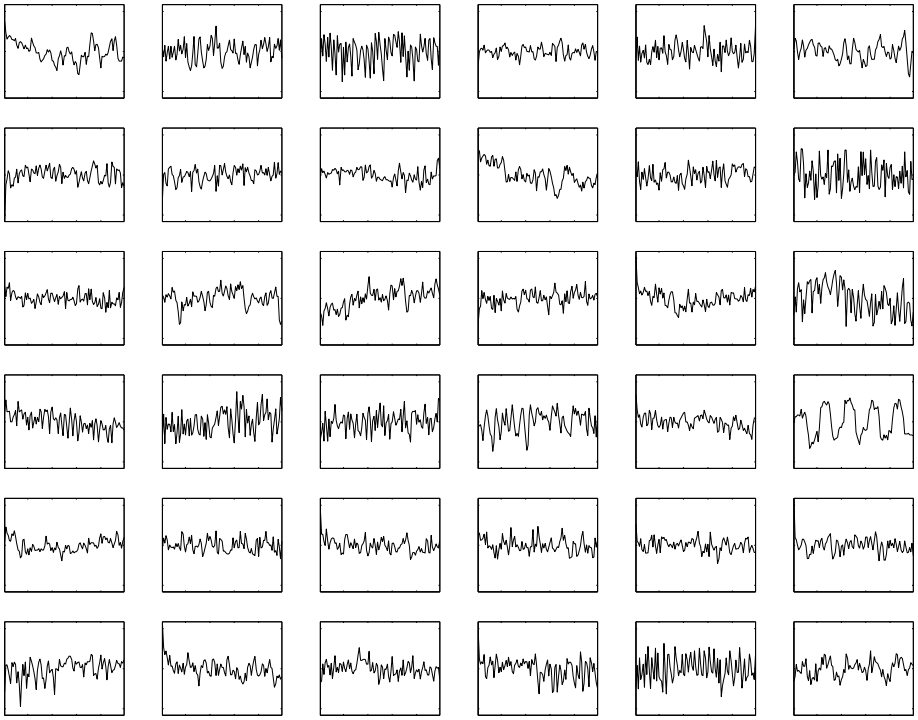


Figure 2: Determined IC time courses for the topographic ICA model.

3.2. Comparison based on Correlation with the Reference Function

An interesting aspect can be observed if we compare the computed reference functions at the maximum correlation for topographic ICA with both the "neural gas" network and the FastICA. Figure 3 visualizes the computed reference functions for all model–free methods. We see that the reference function for the topographic ICA is quite similar to the "neural gas" network. It's important to point out, that the reference function corresponding to topographical ICA has at large the shape of that corresponding to FastICA. However, it contains at small scale the details of the "neural gas" network.

3.3. Comparison based on Activation Maps

Each ICA component map is described by a distribution of values, one per voxel. These values describe the relative amount a given voxel is modulated by the activation of that component. A way to determine and visualize voxels belonging to a particular component map, the map values are scaled to z–scores (the number of standard deviations from the map mean). An active voxel for that component is a voxel whose absolute z–scores are above a predefined threshold(e.g., $|z| > 2$). The z–scores serve for visualization purposes, and have no statistical relevance [MJM+98].

To further differentiate in terms of detecting activation clusters, we take a look at the activation z–maps for all model–free analysis methods as illustrated

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 509-514

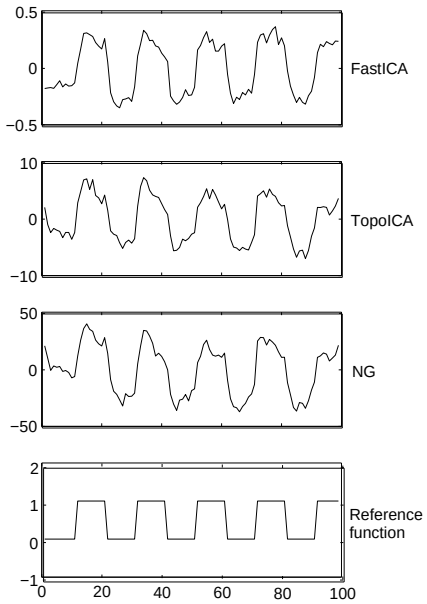


Figure 3: Computed reference functions for the three techniques ("neural gas" network for 36 code vectors, and the topographic ICA and the FastICA for 16 different components.

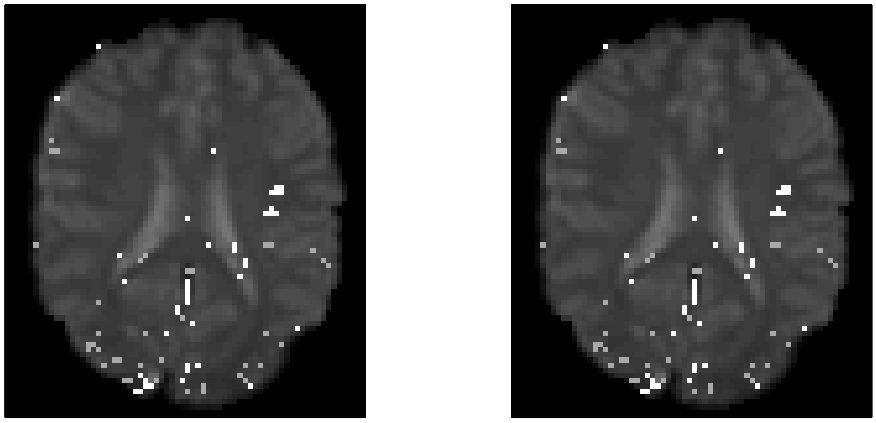
by Figure 4. Here again, it's evident that the topographic ICA unifies the strenghts of the two other techniques.

4. Conclusion

We have adopted a new ICA algorithm for fMRI data analysis, the topographic independent analyisis model. It represents an unifying model that combines topographic mapping with ICA. This new approach is experimentally compared with an already proven ICA algorithm for fMRI, the FastICA, and with an advanced unsupervised clustering method, the "neural gas" network. The goal of our study was to confirm the robustness and reliability in extracting task–related activation maps and time–courses from fMRI data sets. Based on the obtained results, we can conclude that by introducing a topographic order between the components, we are able to identify fine structures from a large noisy data set. While at large (outside the defined grid), the algorithm separates functional response from artifacts, within the grid it classifies time signals according to their temporal similarity. Thus it removes the noise sensitivity of clustering methods, and it differentiates better finer activation structures than transformation–based methods.

The obtained results demonstrate how topographic ICA can provide additional information even in those situations where the obtained results can be foreseen or confirmed by conventional statistical analysis methods. This aspect is of great value for planning experiments under more complex conditions.

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 509-514



(a) (b)

Figure 4: Differences in the activation maps between topographic ICA, "neural gas", and FastICA. (a) FastICA versus TopoICA, and (b) TopoICA versus FastICA. The darker pixels show voxels being active with the first technique, while the lighter ones show only those active with the second technique. The number of code vectors and of components was set equal to eight.

References

[Car98] J. F. Cardoso. *Multidimensional independent componnet analysis. Proc. IEEE ICASSP, Seattle*, 5 1998.

[HH00] A. Hyvarinen und P. Hoyer. *Emergence of Phase- and Shift-Invariant Features by Decomposition of Natural Images into Independent Feature Subspaces. Neural Computation*, 12:1705–1720, 7 2000.

[HH01] A. Hyvarinen und P. Hoyer. *Topographic Independent Component Analysis. Neural Computation*, 13:1527–1558, 7 2001.

[Koh96] T. Kohonen. *Emergence of invariant-feature detectors in the adaptive-subspace self-organizing map. Biological Cybernetics*, 75:281–291, 1 1996.

[MJM+98] M. McKeown, T. Jung, S. Makeig, G. Brown, T. Jung, S. Kindermann, A. Bell und T. Sejnowski. *Analysis of fMRI data by blind separation into independent spatial components. Human Brain Mapping*, 6:160–188, 8 1998.

[WCM92] R. Woods, S. Cherry und J. Mazziotta. *Rapid automated algorithm for aligning and reslicing PET images. Journal of Computer Assisted Tomography*, 16:620–633, 8 1992.

[WLD+02] A. Wismüller, O. Lange, D. Dersch, G. Leinsinger, K. Hahn, B. Pütz und D. Auer. *Cluster Analysis of Biomedical Image Time–Series. International Journal on Computer Vision*, 46:102–128, 2 2002.