Model–Free Functional MRI Analysis Using Transformation–Based Methods

Thomas Dan Otto, Anke Meyer–Bäse,

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

Monica Hurdal, DeWitt Sumners,

Department of Mathematics,
Florida State University, Tallahassee, Florida 32306-4510 USA

Axel Wismüller, and

Institute for Diagnostic Radiology, Ludwig–Maximilians University, Munich 80336 Germany

Dorothee Auer

Max Planck Institute of Psychiatry, Munich Germany

ABSTRACT

This paper presents new model–free fMRI methods based on independent component analysis. Commonly used methods in analyzing fMRI data, such as the Student's t-test and cross correlation analysis, are model-based approaches. Although these methods are easy to implement and are effective in analyzing data with simple paradigms, they are not applicable in situations in which pattern of neural response are complicated and when fMRI response is unknown. In this paper we evaluate and compare three different neural algorithms for estimating spatial ICA on fMRI data: the Infomax approach, the FastICA approach, and a new topographic ICA approach. A comparison of these new methods with principal component analysis and cross correlation analysis is done in a systematic fMRI study determining the spatial and temporal extent of task-related activation. Both topographic ICA and FastICA outperform principal component analysis and Infomax neural network and standard correlation analysis when applied to fMRI studies. The applicability of the new algorithms is demonstrated on experimental data.

Keywords: Topographic ICA, PCA, Infomax, FastICA, fMRI

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.[1–5] According to both theoretical estimations and experimental results,[1,6,7] 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)[8,9] or independent component analysis (ICA),[10,11] transforms original data into high-dimensional vector space to separate functional response and various noise sources from each other. The major problem of PCA is that it only diminishes second-order dependency between each component. ICA can reduce higher order dependency, but is still limited by stationary distribution and the linear mixture assumption.[11]

Email: amb@eng.fsu.edu Paper-location-URL: http://www.dtro.e-technik.th-darmstadt.de/amb/papers

Independent Component Analyses, Wavelets, and Neural Networks, Anthony J. Bell, Mladen V. Wickerhauser, Harold H. Szu, Editors, Proceedings of SPIE Vol. 5102 (2003) © 2003 SPIE · 0277-786X/03/$15.00

The second method, fuzzy clustering analysis (FCA)[12,13] or self-organizing map (SOM),[14,15] attempts to classify time signals of the brain into several patterns according to temporal similarity among these signals. The problems with both methods are that they have difficulty in finding the activation of a small area, they promote sensitivity to noise and initial values, and they have long processing time.

Among the data–driven techniques, ICA has been shown to provide a powerful method for the exploratory analysis of fMRI data.[16,11] ICA is an information theoretic approach which enables to recover underlying signals, or independent components (ICs) from linear data mixtures. Therefore, it is an excellent method to be applied for the spatial localization and temporal characterization of sources of BOLD activation. ICA can be applied to fMRI both temporal[17] or spatial.[11] Spatial ICA has dominated so far in fMRI applications because the spatial dimension is much larger than the temporal dimension in fMRI. However, recent literature results[18] have suggested that temporal and spatial ICA yield similar results for experiments where two predictable task–related components are present.

2. THE ICA ALGORITHMS

Functional organization of the brain is based on two complementary principles, localization and connectionism. Localization means that each visual function is performed mainly by a small set of the cortex. Connectionism, on the other hand, expresses that the brain regions involved in a certain visual cortex function are widely distributed, and thus the brain activity necessary to perform a given task may be the functional integration of activity in distinct brain systems. It is important to stress out that in neurobiology the term "connectionism" is used in a different sense that that used in the neural network terminology.

According to these principles, it was suggested for the first time in[11] that the multifocal brain areas activated by performance of a visual task should be unrelated to the brain areas whose signals are affected by artifacts of physiological nature, head movements, or scanner noise related to fMRI experiments. Every single above mentioned process can be described by one or more spatially–independent components, each associated with a single time course of a voxel and a component map. It is assumed that the component maps, each described by a spatial distribution of fixed values, is representing overlapping, multifocal brain area of statistically dependent fMRI signals. This aspect is visualized in Figure 1. In addition, it is considered that the distributions of the component maps are spatially independent, and in this sense uniquely specified. Mathematically, this means that if $p_k(C_k)$ specifies the probability distribution of the voxel values C_k in the kth component map, then the joint probability distribution of all n components yields:

$$p(C_1, \cdots, C_m) = \Pi_{k=1}^{n} p_k(C_k) \tag{1}$$

where each of the component maps C_k is a vector $(C_{ki}, i = 1, 2, \cdots, M)$, where M gives the number of voxels. Independency is a stronger condition than uncorrelatedness.[19] It was shown in[11] that these maps are indepedent if the active voxels in the maps are sparse and mostly nonoverlapping. Additionally it is assumed that the observed fMRI signals are the superposition of the individual component processes at each voxel. Based on these assumptions, ICA can be applied to fMRI time–series to spatially localize and temporally characterize the sources of BOLD activation.

Different methods for performing ICA decompositions have been proposed which employ different objective functions together with different criteria of optimization of these functions, and it is assumed that they can produce different results.

The aim of this paper is to evaluate and compare three different neural algorithms for estimating spatial ICA on fMRI data: the Infomax approach,[20] the FastICA approach,[21] and a new topographic ICA approach.[22]

2.1. MODELS Of SPATIAL ICA IN fMRI

In the following we will assume that $\mathbf{X}$ is a $T \times M$ matrix of observed voxel time courses (fMRI) signal data matrix), $\mathbf{C}$ is the $N \times M$ random matrix of component map values, and A is a $T \times N$ mixing matrix containing in its columns the associated time–courses of the N components. Furthermore, T corresponds to the number of scans, and M is the number of voxels included in the analysis.

The spatial ICA (sICA) problem is given by the following linear combination model for the data:

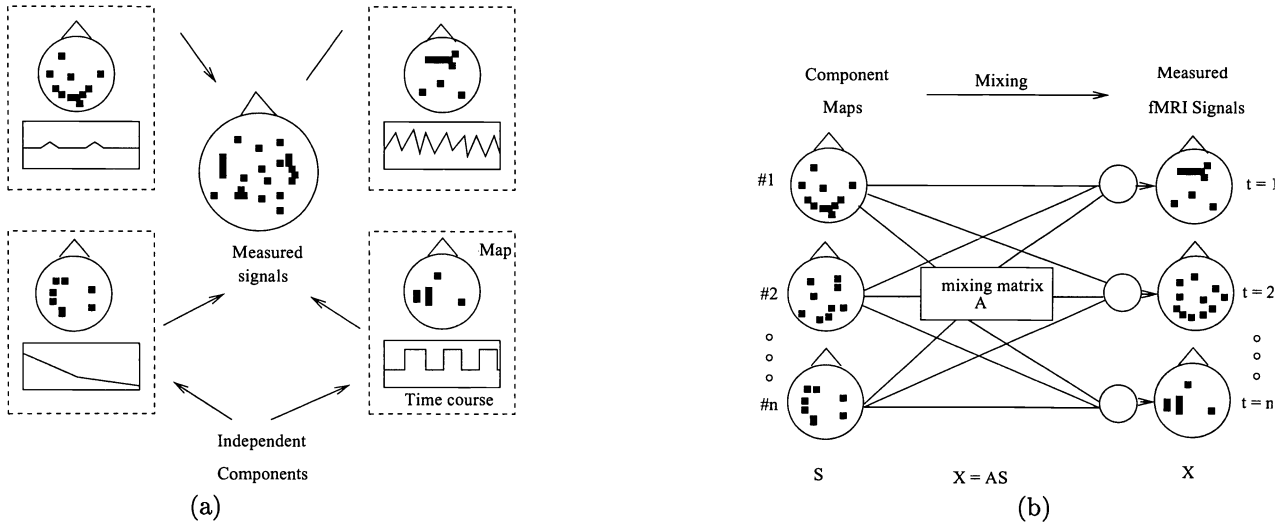

Figure 1. Visualization of ICA applied to fMRI data. (a) Scheme of fMRI data decomposed into independent components, and (b) fMRI data as a mixture of independent components where the mixing matrix **M** specifies the relative contribution of each component at each time point.[11]

$$\mathbf{X} = \mathbf{AC} \tag{2}$$

where no assumptions are made about the mixing matrix **A** and the rows $\mathbf{C_i}$ being mutually statistically independent.

Then the ICA decomposition of **X** can be defined as an invertible transformation:

$$\mathbf{C} = \mathbf{WX} \tag{3}$$

where **W** is an unmixing matrix providing a linear decomposition of data. **A** is the pseudoinverse of **W**.

2.2. INFOMAX APPROACH

The Infomax was the first ICA application to fMRI time–series, and is based on minimization of mutual information.[20] The algorithmic description is given in the following. A self–organizing learning algorithm is described that maximizes the information transferred in a network of linear units. It was shown that the neural network is able to perform ICA estimation and that the nonlinearities in the transfer function introduce higher order statistics. The network architecture is shown in Figure 2.

The network has N input and output neurons, and a $N \times N$ weight matrix **W** connecting the input layer neurons with the output layer neurons. Assuming sigmoidal units, the neurons outputs are given by

$$\mathbf{y} = g(\mathbf{C}), \quad \text{with} \quad \mathbf{C} = \mathbf{WX} \tag{4}$$

where g is a logistic function $g(u_i) = \frac{1}{1+\exp -u_i}$.

The idea of this algorithm is to find an optimal weight matrix **W** such that the output entropy $H(\mathbf{y})$ is maximized. The algorithm initializes **W** to the identity matrix **I**. The elements of **W** are updated based on the following rule

$$\mathbf{W} \leftarrow -\eta(\frac{\partial H(\mathbf{y})}{\partial \mathbf{W}})\mathbf{W}^T\mathbf{W} = -\eta(\mathbf{I} + f(\mathbf{C})\mathbf{C}^T)\mathbf{W} \tag{5}$$

where η is the learning rate. The term $\mathbf{W}^T\mathbf{W}$ in equation (5) was first proposed in,[23] and it avoids matrix inversions and speeds convergence. The vector–function f has the elements

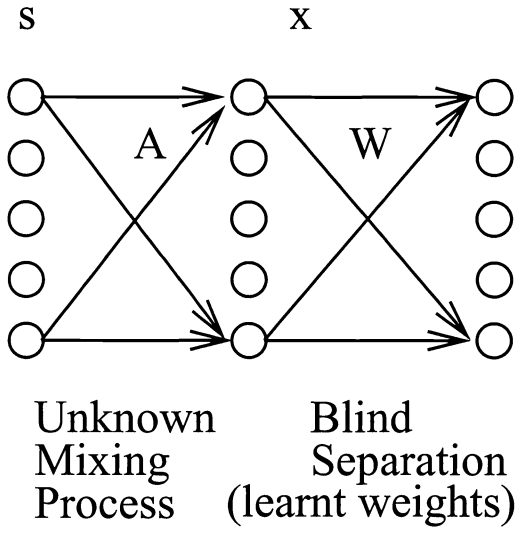

Figure 2. Feedforward neural network for ICA.[20]

$$f_i(u_i) = \frac{\partial}{\partial u_i} \ln g_i'(u_i) = (1 - 2y_i) \tag{6}$$

During training, the learning rate is reduced gradually until the weight matrix $\mathbf{W}$ stops changing appreciably.

Equation (5) represents the so-called "Infomax" algorithm.

The choice of a nonlinearity is determined by the application type. In the context of fMRI, where small activity foci in a large volume are usually expected, the distribution of the target components is assumed to be super–gaussian or sparse. Therefore, a sigmoidal function $g(\dot{})$ is relevant for such an application.

2.3. FastICA APPROACH

The FastICA approach is based on minimization of mutual information but using the negentropy as a measure of non–gaussianity. This approach is both symmetric and hierarchical, and is based on Fixed–Point iterations. To apply this ICA approach, the data must be preprocessed by centering and whitening. We first look at the one–unit version. The unit can be an artificial neuron having a weight vector $\mathbf{w}$ that is updated based on a learning algorithm. It finds a vector $\mathbf{w}$ such that the projection $\mathbf{w}^T\mathbf{x}$ maximizes nongaussianity. As a measure the approximation of the negentropy can be choosen.[24]

The Fast ICA algorithm is a fixed–point iteration scheme for finding a maximum of the nongaussianity of $\mathbf{w}^T\mathbf{x}$. Let g be the derivative of the non–quadratic function. Choices of such nonquadratic functions are either $g(u) = \frac{1}{a_1} \log \cosh a_1 u$ or $g(u) = -\exp -u^2/2$ where $1 \leq a_1 \leq 2$ is a suitable constant.

To estimate several independent components, the one–unit Fast ICA is employed using several units (neurons) with weight vectors $\mathbf{w}, \cdots, \mathbf{w}_n$. To prevent different vectors from converging to the same maxima the outputs $\mathbf{w}_1^T\mathbf{x}, \cdots, \mathbf{w}_n^T\mathbf{x}$ have to be decorrelated after every iteration. For a whitened $\mathbf{x}$ this is equivalent to orthogonalization. There are several known methods to achieve this.[24] Here, only the symmetric decorrelation is considered. It has several advantages over other methods: (1) the weight vectors $\mathbf{w}_i$ are estimated in parallel and not one by one and (2) it doesn't perpetuate the errors from one weight vectors to the next.

The symmetric orthogonalization of $\mathbf{W}$ can be accomplished by involving matrix square roots

$$\mathbf{W} = (\mathbf{W}\mathbf{W}^T)^{\frac{-1}{2}}\mathbf{W} \tag{7}$$

Numerical simplifications of the above equation are given in.[24]

An algorithmic description of the Fast ICA algorithm for estimating several independent components is given below:

1. **Preprocessing:** Whiten the data to give $\mathbf{x}$.
2. **Number of independent components:** Choose m the number of independent components to estimate.
3. **Initialization:** Initialize each weight vector $\mathbf{w}_i, i = 1, \cdots, m$ of unit norm with random numbers.
4. **Adaptation:** Compute the change of every weight vector $\mathbf{w}_i, i = 1, \cdots, m$ accordingly to

$$\mathbf{w}_i \leftarrow E\{\mathbf{x}g(\mathbf{w}_i^T\mathbf{x})\} - E\{g^{'}(\mathbf{w}_i^T\mathbf{x})\}\mathbf{w}_i \tag{8}$$

   where g is defined as $g(u) = \tanh a_1 u$, with $1 \leq a_1 \leq 2$ being a constant.
5. **Symmmetric orthogonalization:** Do a symmetric othogonalization of the matrix $\mathbf{W} = (\mathbf{w}_1, \cdots, \mathbf{w}_m)^T$ by

$$\mathbf{W} \leftarrow (\mathbf{W}\mathbf{W}^T)^{\frac{-1}{2}}\mathbf{W} \tag{9}$$

6. **Continuation:** If the weight matrix $\mathbf{W}$ is not converged, go back to step 4.

The main difference between Infomax and FastICA lies in the updating rule: It is adaptive for the Infomax depending on a learning rate, and it is nonadaptive for FastICA.

2.4. TOPOGRAPHIC ICA APPROACH

Topographic ICA represents a generative model which combines topographic mapping with ICA. As in all topographic mappings, the distance in the representation space given by the topographic grid is related to the distance of the represented components. This distance is defined for topographic ICA by the mutual information implied by higher–order correlations.[22] Thus, a natural distance measure is given in the context of ICA. Traditional topographic mapping methods define distance either based on the Euclidian distance or correlation. The ICA distance measure enables the definition of a topography even if the Euclidian distances are all equal as it is the case with an orthogonal vector space.

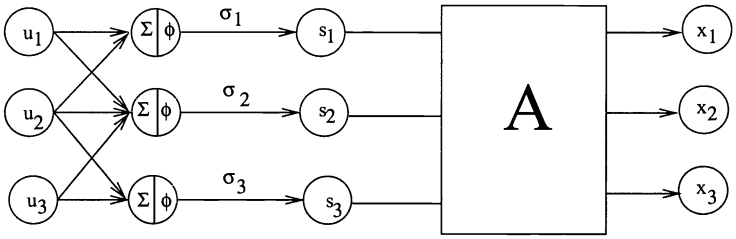

Figure 3. Topographic ICA model.[22] The variance generated variables u_i are randomly generated, and mixed linearly inside their topographic neighborhoods. This forms the input to nonlinearity ϕ, thus giving the local variance σ_i. Components s_i are generated with variances σ_i. The observed variables are x_i are obtained as with standard ICA from the linear mixture of the components s_i.

In the generative model described in[22] and shown in Figure 3, s_i are independent given their variances σ_i^2. Dependence among the s_i is enforced by the variance dependence. Obeying the principle of topography, the variances of only neighboring components are positively correlated, while the others are indepedent. By using a neighborhood function $h(i,j)$, the variance σ_i is given by:

$$\sigma_i = \phi\left(\sum_{k=1}^{n} h(i,k)u_k\right) \tag{10}$$

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. The components s_i are given by the following relationship:

$$s_i = z_i \sigma_i \tag{11}$$

where z_i is a random variable having the same distribution as s_i while σ_i^2 is fixed to unity. u_i and z_i are mutually independent.

The most important properties of the topographic ICA are: (1) all the components are uncorrelated, (2) components far from each other are independet, (3) neighboring components have positively correlated energies s_i^2 and s_j^2. The classic ICA results from the topographic ICA by setting $h(i,j) = \delta_{ij}$.

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 update rule for the weight vector $\mathbf{w}_i$ is derived from a gradient algorithm:

$$\Delta \mathbf{w}_i \propto E\{\mathbf{x}(\mathbf{w}_i^T \mathbf{x}) r_i)\} \tag{12}$$

where

$$r_i = \sum_{k=1}^{n} h(i,k) g(\sum_{j=1}^{n} k(k,j)(\mathbf{w}_j^T \mathbf{x})^2) \tag{13}$$

The function g is the derivative of $G = -\alpha_1 \sqrt{u} + \beta_1$. After every iteration, the vectors $\mathbf{w}_i$ in equation (12) 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.

It's useful to point out some differences between topographic ICA and other topographic mappings: (1) topographic ICA finds a decomposition into independent components, while topographic mappings find cluster centers or codevectors, and (2) the similarity of two vectors in topograpic ICA is based on higher–order correlations and not defined by an Euclidian distance or dot–product. However, if the data is prewhitened, the dot–product in the data space is equivalent to to correlation in the original space.[22]

3. RESULTS AND DISCUSSION

FMRI data were recorded from six subjects (3 female, 3 male, age 20-37) performing a visual (welcher?) 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. The first scan was discarded for remaining saturation effects. Motion artifacts were compensated by automatic image alignment (AIR,[25]).

3.1. ESTIMATION OF THE ICA MODEL

To decide to what extent spatial ICA of fMRI time–series depends on the employed algorithm, we have first to look at the optimal number of principal components selected by PCA and used in the ICA decomposition. ICA is a generalization of PCA. In case no ICA is performed, then the number of independent components equals zero, and this means there is no PCA decomposition performed.

Therefore, it is significant to find a fixed number of ICs that can theoretically predict new observations in same conditions, assuming the basic ICA model actually holds. To do so, we compared the three proposed algorithms, and additionally PCA for 8, 16, and 36 components in terms of: (1) the component time course found by each of the four techniques most closely matching the visual task reference function, (2) the normalized quantization error, and (3) the necessary simulation time.

The obtained results are plotted in Figure 4. We distinguish between preprocessed data (whitening and sphering) and unprocessed data. In the following we will give the set parameters. For PCA, no parameters had to be set.

For Infomax we choose: (1) $\epsilon = 10^{-6}$, and (2) 10^5 as the maximal number of iterations. For FastICA we choose: (1) $\epsilon = 10^{-6}$, (2) 10^5 as the maximal number of iterations, and (3) the nonlinearity $g(u) = \tanh u$. And last, for topographic ICA we set: (1) stop criterium is fullfilled if the synaptic weights difference between two consecutive iterations is less than $10^{-5} \times$ Number of IC, (2) the function $g(u) = u$, and (3) 10^4 is the maximal number of iterations.

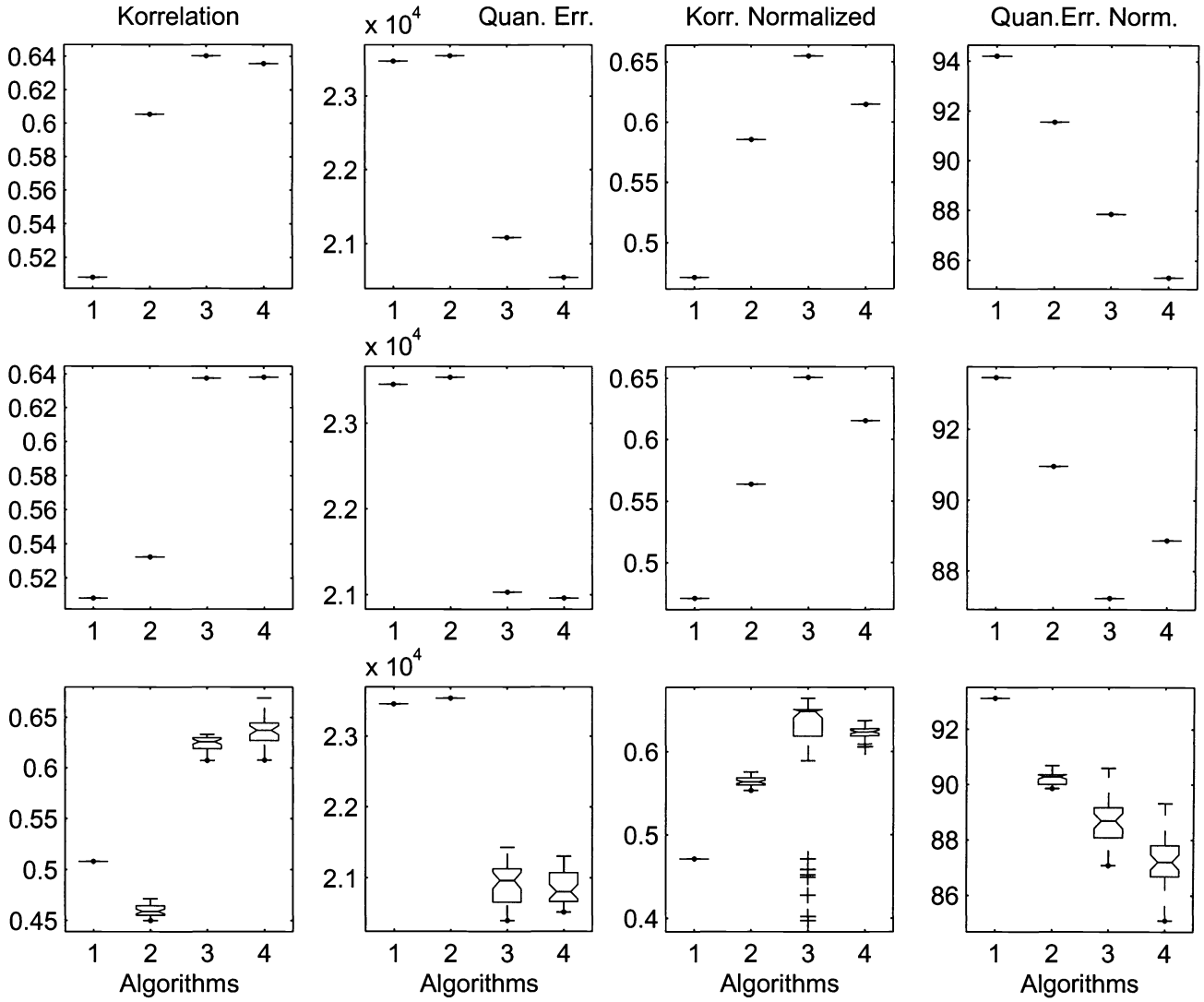

Figure 4. Estimation results for the optimal independent components for the four techniques for normalized and unnormalized data, and considering 40 runs. The techniques 1 to 4 are: PCA, Infomax, FastICA, and topographic ICA. The plots show the correlation coefficients between the found CTR components to the given task reference function, and the normalized quantization error. The number of independent components was equal to 8 for the first row, equal to 16 for the second, and equal to 32 for then last row.

The necessary simulation time for the four mentioned techniques in terms of the selected number of components is given in Figure 5.

A close scrutinity at both Figures 4 and 5 reveals that the best fixed number of ICs is at 8 and 16, more than this number will result in encoding noise. And also, the quantization error indicates that 8 and 16 ICs encode most of the real (non–artifact) components. It can be also seen that almost all ICA techniques operate more reliable at a lower number of components. The following plots and evaluations are based on this found optimal number of ICs.

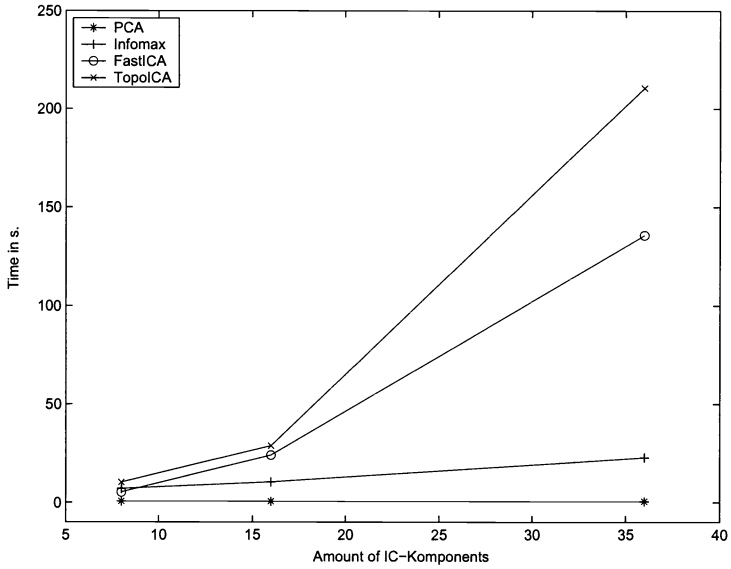

Figure 5. Necessary simulation time for the four techniques, PCA, Infomax, FastICA, and topographic ICA, in terms of selected components.

3.2. CHARACTERIZATION OF TASK–RELATED EFFECTS

For all subjects, and runs, unique task–related activation maps and associated time–courses were obtained by all three ICA techniques.

Figure 6 shows for 8 ICs the component time course most closely associated with the visual task for all three ICA techniques and PCA. The best results are achieved by FastICA and topographic ICA, yielding almost identical reference functions, and a correlation coefficient of $r = 0.92$ between those two. PCA has a much lower correlation and also Infomax.

However, the activation maps do not reveal much of a difference between the ICA techniques. Figure 7 shows the activation maps as a comparison of results obtained by the four techniques.

The found results in Figure 7 motivate a closer look at the activation maps of each technique such that a more general statement regarding the differentiation capabilities of each technique is possible. Two representation modalities are choosen: one based on the so–called z–statistics, and the other based on the correlation value. Figure 8 shows the activation maps for each of the four techniques based on the z–statistics. The lighter a pixel is, the more correlated is the corresponding voxel to the CTR. The FastICA and TOPOICA activation maps appear very similar to each other. FastICA and topographic ICA show additional active voxels which appear based on the intensity to be real active voxels and not noise.

Figure 9 shows the activation map based on correlation representation. The results differ significantly from those of the activation z–maps.

4. CONCLUSION

In the present paper, we have experimentally compared two standard ICA algorithms already adopted in the fMRI literature with a new algorithm, the topographic ICA. The goal of the paper was to determine the robustness and reliability of extracting task–related activation maps and time–courses from fMRI data sets. The success of ICA methods is based on the condition that the spatial distribution of brain areas activated by task performance must be spatially independent of the distributions of areas affected by artifacts.

It has been shown that both topographic ICA and FastICA outperform in terms of found component time course, quantization error, and activartion maps principal component analysis and Infomax neural network when applied to fMRI studies. The applicability of the new algorithms is demonstrated on experimental data.

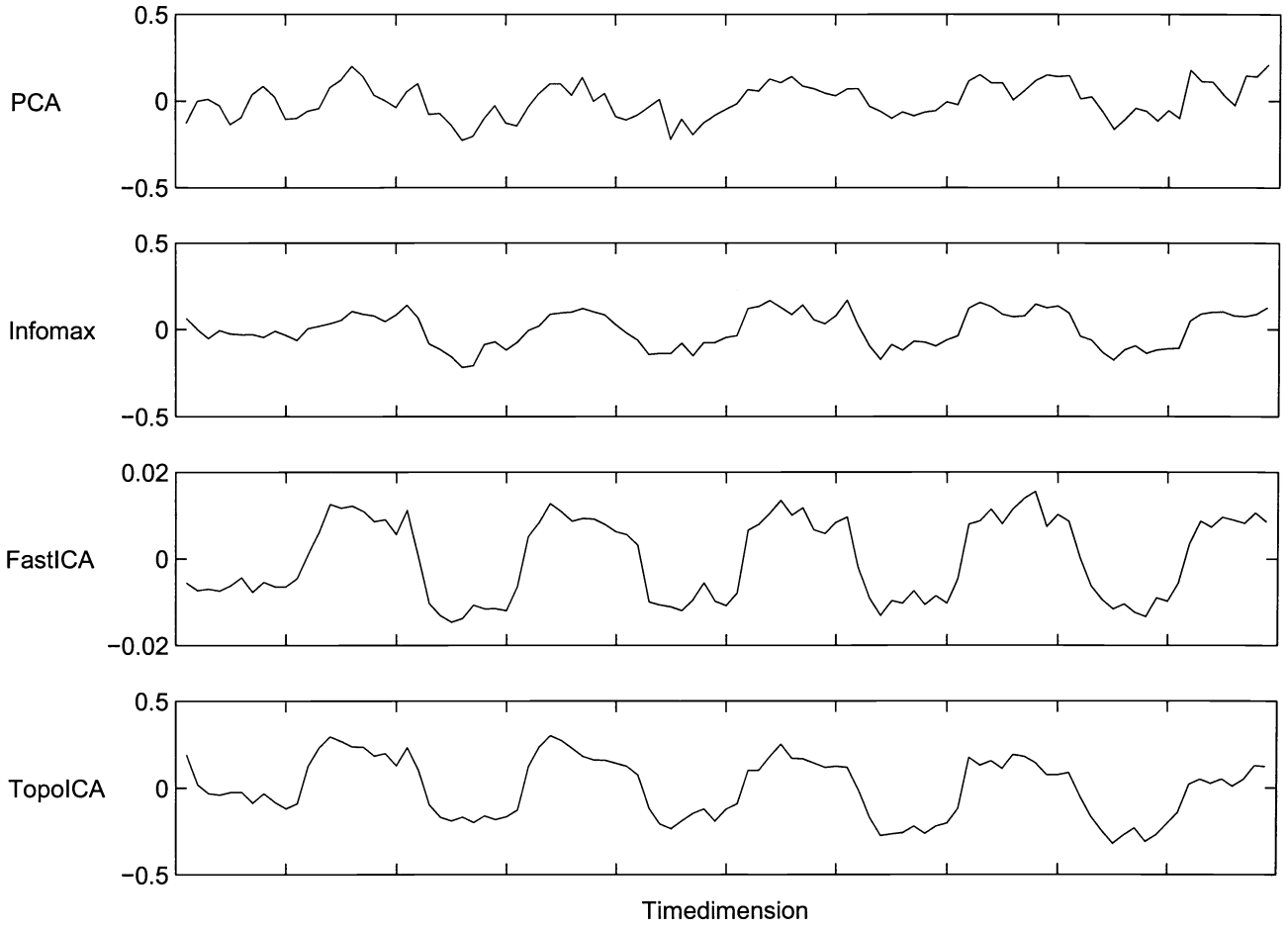

Figure 6. Computed reference functions for the four techniques, PCA, Infomax, FastICA, and topographic ICA, for 16 ICs. The found correlation coefficients are: $r = 0.69$ for PCA, $r = 0.85$ for Infomax, $r = 0.94$ for FastICA, and $r = 0.85$ for topographic ICA.

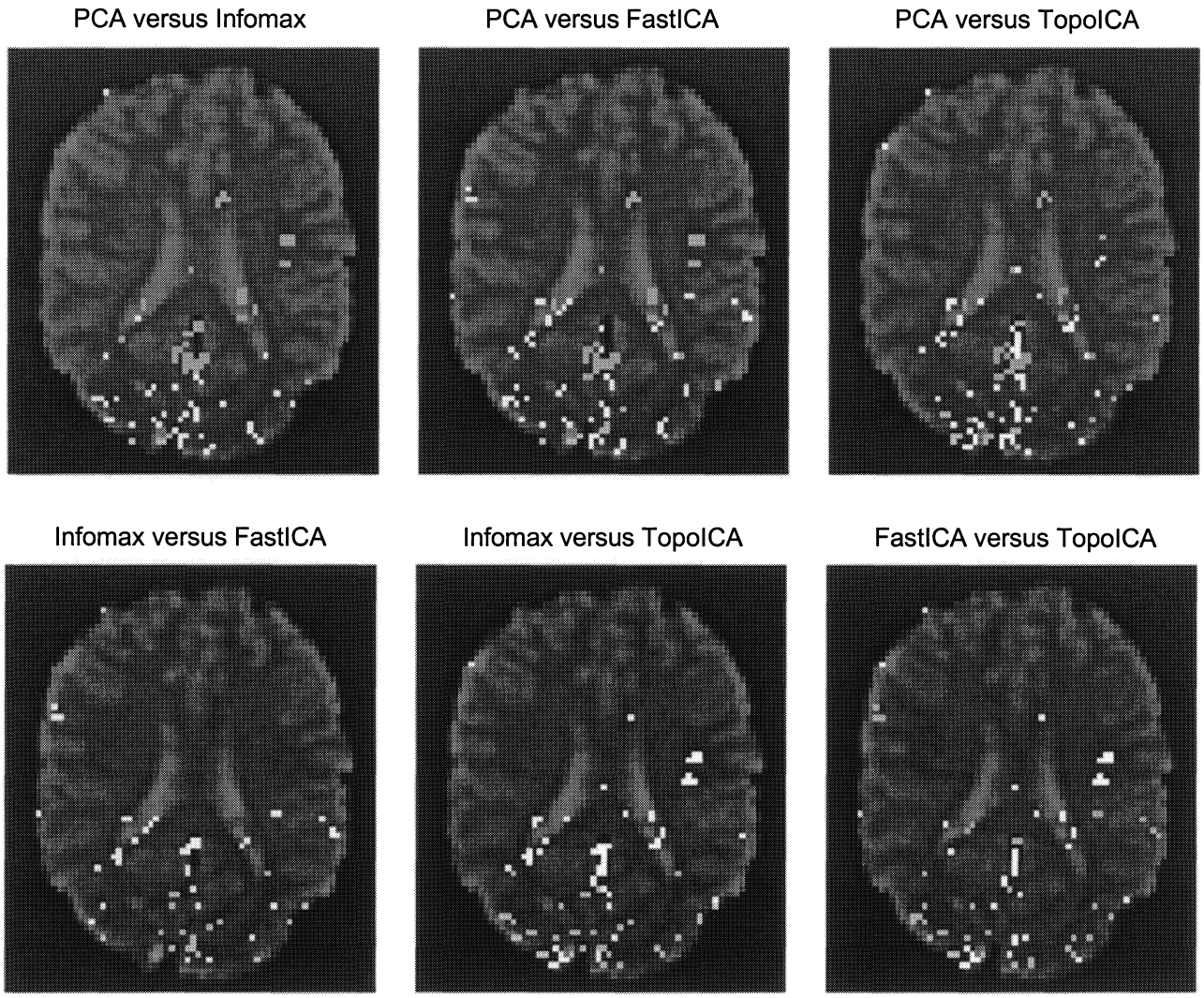

Figure 7. Comparison of the results for the four techniques, PCA, Infomax, FastICA, and topographic ICA. The shown activation maps show only the pairwise differences. The darker pixels show active voxels for the first technique, while the lighter pixels show the active voxels for the second technique.

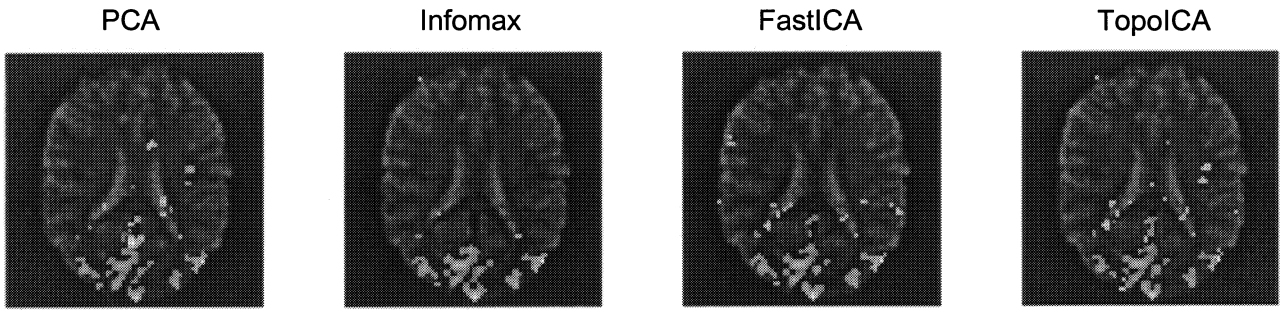

Figure 8. Activation z–maps for the four techniques, PCA, Infomax, FastICA, and topographic ICA. The threshold for an active pixel was set at $|z| > 2$. The maximum $|z|$–values for each technique are: 9.0 (PCA), 9.9 (Infomax), 9.2 (FastICA), and 8.6 (topographic ICA).

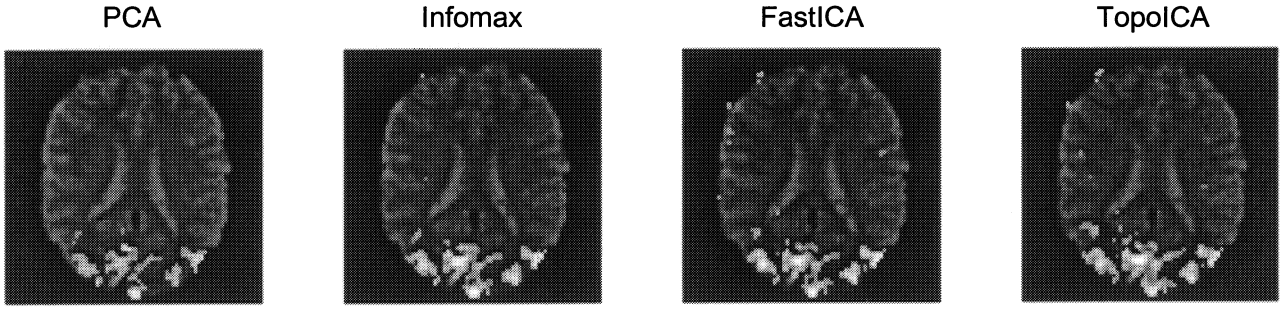

Figure 9. Activation maps based on correlation for the four techniques, PCA, Infomax, FastICA, and topographic ICA. A voxel is considered active by correlation if $r \geq 0.47$. The maximal correlation coefficient r obtained by each of the four techniques is: 0.83 (PCA), 0.88 (Infomax), 0.91 (FastICA), and 0.94 (topographic ICA).

REFERENCES

1. K. Kwong, J. Belliveau, and D. Chesler, "Dynamic magnetic resonance imaging of human brain activity during primary sensor stimulation," 8 1992.
2. S. Ogawa, D. Tank, and R. Menon, "Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging," 8 1992.
3. P. Bandettini, E. Wong, R. Hinks, R. Tikofski, and J. Hyde, "Time course epi of human brain function during task activation," 8 1992.
4. J. Frahm, K. Merboldt, and W. Hanicke, "Functional mri of human brain activation at high spatial resolution," 8 1992.
5. K. Kwong, "Functinal magnetic resonance imaging with echo planar imaging," 8 1992.
6. J. Boxerman, P. Bandettini, K. Kwong, and J. Baker, "The intravsacular contribution to fmri signal change: Monte carlo modeling and diffusion-weighted studies in vivo," 8 1995.
7. S. Ogawa, T. Lee, and B. Barrere, "The sensitivity of magnetic resonance image signals of a rat brain to changes in the cerebral venous blood oxygenation activation," 8 1993.
8. J. Sychra, P. Bandettini, N. Bhattacharya, and Q. Lin, "Synthetic images by subspace transforms i. principal components images and related filters," 8 1994.
9. W. Backfrieder, R. Baumgartner, M. Samal, E. Moser, and H. Bergmann, "Quantification of intensity variations in functional mr images using rotated principal components," 8 1996.
10. M. McKeown, T. Jung, S. Makeig, G. Brown, T. Jung, S. Kindermann, A. Bell, and T. Sejnowski, "Spatially independent activity patterns in functional magnetic resonance imaging data during the stroop color-naming task," 8 1998.
11. M. McKeown, T. Jung, S. Makeig, G. Brown, T. Jung, S. Kindermann, A. Bell, and T. Sejnowski, "Analysis of fmri data by blind separation into independent spatial components," 8 1998.
12. G. Scarth, M. McIntrye, B. Wowk, and R. Samorjai, "Detection novelty in functional imaging using fuzzy clustering," 8 1995.
13. K. Chuang, M. Chiu, C. Lin, and J. Chen, "Model-free functional mri analysis using kohonen clustering neural network and fuzzy c-means," 8 1999.
14. H. Fisher and J. Hennig, "Clustering of functional mr data," 8 1996.
15. S. Ngan and X. Hu, "Analysis of fmri imaging data using self-organizing mapping with spatial connectivity," 8 1999.
16. K. Arfanakis, D. Cordes, V. Haughton, C. Moritz, M. Quigley, and M. Meyerand, "Combining independent component analysis and correlation analysis to probe interregional connectivity in fmri task activation datasets," 8 2000.
17. B. Biswal and J. Ulmer, "Blind source separation of multiple signal sources of fmri data sets using independent component analysis," 8 1999.
18. V. Calhoun, T. Adali, G. Pearlson, and J. Pekar, "Spatial and temporal independent component analysis of functional mri data containing a pair of task–related waveforms," 8 2001.
19. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1986.
20. A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," 12 1995.
21. A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," 5 1999.
22. A. Hyvarinen and P. Hoyer, "Topographic independent component analysis," 7 2001.
23. S. Amari, A. Cichocki, and H. H. Yang, "A new learning algorithm for blind signal separation," 12 1996.
24. A. Hyvarinen and E. Oja, "Independent component analysis: algorithms and applications," 5 2000.
25. R. Woods, S. Cherry, and J. Mazziotta, "Rapid automated algorithm for aligning and reslicing pet images," 8 1992.