

INDEPENDENT COMPONENT ANALYSIS APPLIED TO fMRI DATA: A GENERATIVE MODEL FOR VALIDATING RESULTS

V. Calhoun^{1,2}, T. Adali,² and G. Pearlson¹

¹Johns Hopkins University Division of Psychiatric Neuro-Imaging,
600 N. Wolfe St., Baltimore, MD 21205

²University of Maryland Baltimore County, Dept. of CSEE, Baltimore, MD 21250

ABSTRACT

We introduce and apply a synthesis/analysis model for analyzing functional Magnetic Resonance Imaging (fMRI) data using independent component analysis (ICA). Our model assumes statistically independent spatial sources in the brain. We also assume that the fMRI scanner acquires overdetermined data such that there are more time points than brain sources. We discuss the properties of each of the signals present in the model. The analysis portion of the model includes several candidates for spatial smoothing, ICA algorithm, and data reduction. We use the Kullback-Leibler divergence between the estimated source distributions and the “true” distributions as a measure of the optimality of the final ICA decomposition. Using this model, we generate fMRI-like data and optimize the analysis stage as a function of ICA algorithm, data reduction scheme, and spatial smoothing.

1 INTRODUCTION

Functional MRI (fMRI) is a tool which measures changes in blood flow and blood oxygenation. When brain neurons are activated, there is a resultant localized change in blood flow and oxygenation. These hemodynamic changes are delayed by several seconds following the electrical activity in the brain [1]. fMRI analysis approaches range from model-based to exploratory, although model-based approaches are by far the most utilized [2]. Model-based approaches are extremely useful when the time course of the hemodynamic response can be inferred *a priori*, however the hemodynamics of the brain are still being studied and good *a priori* assumptions are sometimes not ascertainable. The recent application of independent component analysis (ICA) to fMRI data has proven a useful approach [4]. When applied to fMRI, ICA can be used to separate sources that are independent in either time or space [5]. For clarity, the following discussion will be limited to spatial-ICA although extension to temporal-ICA is straightforward.

We apply spatial ICA by assuming that during a given fMRI experiment there are a number of brain regions (networks) which are spatially independent from one another (sources) and are *mixed* together via a network specific hemodynamic time course. This is a reasonable assumption given that (a) brain networks tend to be distributed differently and (b) the changes in blood oxygenation due to neuronal firing often cover large areas (due to draining veins) and appear linear to first approximation at least within sources [3].

Two important steps in the analysis stage include data reduction and spatial smoothing. Because of the high dimensionality of fMRI data, a data reduction scheme is typically applied prior to ICA. The dimensionality of the data in fMRI (how many image volumes are acquired) is determined by the repeat time (TR) parameter. This can be changed from scan to scan and has no relationship to the number of sources in the brain. We assume that more time points are acquired than the number of brain sources, an assumption justified for many fMRI experiments. In addition to data reduction, a spatial smoothing filter is typically applied to fMRI data in model-based approaches. This can also be utilized for an ICA analysis and may provide better results due to the smoothness of the hemodynamic sources.

ICA has only recently been applied to fMRI data [4]. It has proved promising, but there is a need to study the properties of ICA as applied to fMRI data. There is also a need to develop ICA methods to address specific issues involved in fMRI. Our work is to our knowledge the first attempt to generate realistic fMRI data and optimize all stages of the ICA processing.

The model we introduce (see Figure 1) provides a structured approach for the problem and enables us to generate more realistic fMRI-like sources for validating the ICA decomposition. Our goal is to understand the interaction of the assumptions made at all stages including a) preprocessing, b) data reduction (clustering and principal component analysis (PCA)), and c) ICA estimation (in particular using infomax [6], maximum likelihood [7], negentropy [8]). Using this model, we generate sources similar to those encountered in fMRI data. We then generate fMRI-like data and optimize the analysis as a function of ICA algorithm, data reduction scheme, and spatial smoothing.

2 THE MODEL

We introduce the following model to provide a framework for understanding ICA as applied to fMRI data (see Figure 1). We assume a generative model for the data including the brain and the fMRI scanner, and discuss the properties of the signals from each. Our model has two main parts, data generation and data processing.

2.1 Data Generation (Synthesis)

In the *data generation block* we assume that there are a set of statistically independent hemodynamic source locations in the brain (indicated by $s_i(v)$ at location v for the i^{th} source). The sources

$$\mathbf{s}(v) \triangleq [s_1(v), s_2(v), \dots, s_N(v)]^T \quad (1)$$

have weights that specify the contribution of each source to each voxel; these weights are multiplied by each source's hemodynamic time course. These sources, $\mathbf{s}(v)$, are assumed to be spatially independent, that is, the joint source distribution can be represented as

$$p_{\mathbf{s}}(\mathbf{s}) \triangleq p_{s_1, s_2, \dots, s_N}(s_1, s_2, \dots, s_N) = \prod_{i=1}^N p_{s_i}(s_i) \quad (2)$$

This is a reasonable assumption which imposes the condition that a different set of areas (perhaps overlapping) are involved for each signal. The weights of the hemodynamic sources at each voxel location v serve as the inputs to our model. Additional assumptions are that each source $s_i(v)$ is spatially correlated (smooth) in v due to the vascular point spread function that induces local correlations [9] and that the vector $\mathbf{s}(v)$ consists of signals of interest that are not normally distributed [10]. The signal $\mathbf{s}(v)$ is a continuous valued function and continuous in v .

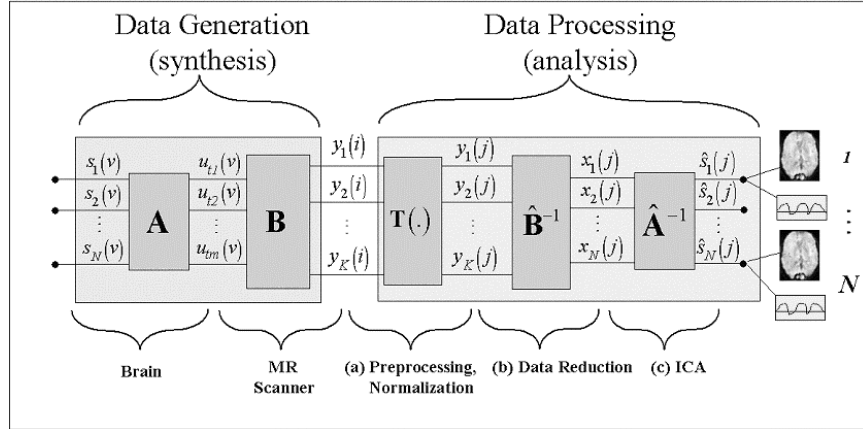


Figure 1: Model for fMRI data in which spatially independent sources are mixed linearly by the hemodynamic response and then overdetermined by the data acquired by the fMRI scanner.

It is assumed that each of the N sources are added together so that a given voxel contains a mixture of the sources, each of which fluctuates according to its weighted

hemodynamic time course. This *linear mixing* is represented by the system, \mathbf{A} , and yields

$$\mathbf{u}(v) \triangleq [u_{t1}(v), u_{t2}(v), \dots, u_{tm}(v)]^T \quad (3)$$

where $t1, \dots, tm$ indicate ideal samples of the mixed source signal (although at this point in the model we haven't imposed any particular sampling scheme). The elements of the vector $\mathbf{u}(v)$ are not independent of one another. This seems reasonable if one considers that the vascular signals are mixed linearly because they are (a) generated from overlapping neural sources and (b) further spread out in space from the original neural sources due to the vasculature [9]. Additionally, the mixtures of the spatial source observations may be temporally correlated due to the hemodynamic response function [11]. The signal $\mathbf{u}(v)$ is a continuous valued function, continuous in v .

The first portion of the data generation block takes place within the brain. The second portion of the data generation block involves the fMRI scanner at which point we assume K discrete time points are acquired, more than the number of sources in the brain. During the fMRI acquisition, the data is discretely sampled in space (at locations indicated by $i = 1, 2, \dots, M$ where M is the number of voxels), after which the data is preprocessed (motion corrected and smoothed) and spatially normalized into a standard space (at location indicated by $j = 1, 2, \dots, M$). The sampling of the brain's hemodynamics with the fMRI scanner results in

$$\mathbf{y}(i) \triangleq [y_1(i), y_2(i), \dots, y_N(i)]^T. \quad (4)$$

The signal $\mathbf{y}(i)$ may be temporally correlated due to magnetic equilibrium effects. We know that $\mathbf{y}(i)$ is spatially correlated due to the reconstruction algorithm (zero-filled FFT) as well as to the point-spread function of the scanner.

2.2 Data Processing (Analysis)

In the *data processing block* we have a transformation $\mathbf{T}(\cdot)$ representing a number of possible preprocessing stages, including slice phase correction, motion correction, spatial normalization and smoothing. The matrices $\hat{\mathbf{B}}^{-1}$ and $\hat{\mathbf{A}}^{-1}$ indicate systems designed to find signal estimates $\hat{\mathbf{x}}(j)$ and $\hat{\mathbf{s}}(j)$, respectively. The resultant source estimated $\hat{\mathbf{s}}(j)$, along with the unmixing matrix $\hat{\mathbf{A}}^{-1}$, can then be presented as fMRI activation images and fMRI time courses, respectively.

We have incorporated a data reduction stage into our model as it is a typical step employed due to the high dimensionality of fMRI data. Additionally, if one does not perform data reduction, the interpretation of the component maps can be a very time consuming task and a large majority of these maps are not useful (non-interesting). The goal is to reduce the data and estimate fewer components with the information in the non-interesting components either being 1) discarded, or 2) redistributed into

a new set of non-interesting components. Data reduction aids both in decreasing computational load and decreasing data interpretation time.

We discuss two approaches for data reduction in this paper, principal component analysis (PCA) and clustering. PCA (the most commonly used approach) projects the data onto a lower dimensional space while minimizing the sum of the square of the error between the original data set and the reduced data set. It is important to note that PCA utilizes a second order criterion to determine the reduction space while clustering does not. In [13], we demonstrate a case where the emphasis of second order statistics in the case of PCA can produce sub optimal results.

The final stage in our model is the unmixing (ICA) stage. Two well-known ICA approaches are infomax and negentropy. The infomax algorithm [6] minimizes the mutual information that the output of a neural network has about its input. The fastICA implementation of Hyvarinen maximizes the negentropy (approximated by kurtosis) of the output sources [8]. Both of these ICA algorithms are implemented in Matlab code available on the Internet and both have been used for processing fMRI data [4,5].

3 METHODS

We demonstrate the application of our analysis/synthesis model in a simulated experiment. Incorporating the information we have discussed regarding fMRI data, we generate a representative set of fMRI-like signals to create a realistic simulated data set. We then optimize the parameters discussed above (smoothing, data reduction, and the ICA method used) for this data set and discuss the results.

3.1 Generating fMRI-“like” signals

There are several types of signals that can be encoded within the hemodynamic signals measured by fMRI. In order to generate realistic simulated signals, we describe the types of signals found in fMRI data. *Task-related* signals are those that most strongly resemble the activation paradigm used in the fMRI experiment. Such paradigms are usually periodic and often a square-wave is used as a first approximation of the expected time course. *Transiently task-related* signals tend to be periodic as well, but are shorter in duration than the task-related signals. *Motion-related* signals can also be present and tend to be changes across large regions of the image (particularly at the edges of images). They can be slowly varying with time with perhaps a large transient change. The *function-related* signal manifests as similarities between voxels within a particular functional domain (e.g. the motor cortex on one side of the brain will correlate most highly with voxels in the motor cortex on the opposite side of the brain) [12]. The temporal fluctuations of this signal typically look random. A summary of the signals discussed is presented in the table below.

Table 1: Summary of Signals

<i>Temporal Pattern</i>	<i>Signal Types, $s(v)$</i>
Periodic, slowly varying	Task-related
Periodic, transient	Transiently task-related
Random fluctuations	Function-related
Very slowly varying with large transients	Motion-related

3.2 Evaluation

With the analysis/synthesis model, we introduced, we can evaluate performance of analysis methods by comparing the estimated sources with the true distributions. We use the relative entropy (Kullback-Leibler (KL) divergence) between the estimated and true distributions as a measure of the goodness of the estimation. This can be represented as:

$$D(\mathbf{s} \parallel \hat{\mathbf{s}}) = \int p_s(\mathbf{y}) \ln \left(\frac{p_s(\mathbf{y})}{\hat{p}_s(\mathbf{y})} \right) d\mathbf{y}, \quad (5)$$

where $p_s(\mathbf{y})$ and $\hat{p}_s(\mathbf{y})$ are the joint distributions of the true and estimated sources, respectively. Since the sources and estimated sources are independent (Equation 2), the joint distributions can be factorized and $D(\mathbf{s} \parallel \hat{\mathbf{s}})$ can be written as the sum of the marginal divergences (provided the true and estimated sources are ordered in the same way).

3.3 Experiment

Three stages in the analysis are varied in our optimization: smoothing, data reduction, and the ICA method used. We compare three Gaussian smoothing kernels: the full width at half maximum of 0 (no smoothing), 2 voxels, or 4 voxels. Two data reduction methods, 1) principal component analysis (PCA) and 2) clustering are compared along with two ICA methods, 1) infomax and 2) fastICA. We use the k-means algorithm to perform clustering. An additional preprocessing step that is often utilized is to pre-whiten the mixed sources. We apply pre-whitening prior to the ICA estimation stage regardless of which data reduction stage is utilized.

We generate sources similar to those described in Table 1. Three of the sources were simulated fMRI activations including task-related (source 1), transiently task-related (source 2) with both positive and a negative activations, and function-related signals (source 4). Note that the activations are smooth (according to this property of the sources described in section 2.1). The time courses of source 1 and source 2 are temporally smooth also according to the properties described in section 2.1. Source 3 simulates motion resulting in bright voxels on one side and dark voxels on

the opposite side. The motion is simulated as slow sinusoidal rotation out of plane with a large transient shift representing sudden movement.

Once the data set was generated, noise was added. Several types of noise were added: 1) temporally correlated Gaussian noise, 2) uncorrelated Gaussian noise, and 3) realistic fMRI scanner noise. We added temporally correlated noise because there is a well-known temporal correlation present in the noise in fMRI data. The fMRI noise was extracted from an fMRI experiment. The voxels occurring outside the head (and expected to include scanner noise, but not to include signal from the brain) were added to the simulated data set. The two Gaussian noise sources were utilized to enable perturbation of the noise across trials (the fMRI noise was not sampled from a distribution but was added to increase the complexity of the noise). We looked at results for two signal-to-noise (SNR) levels. Since the sources have different amplitudes, we use the standard deviation over all the sources and divide this value by the standard deviation over all the noise sources as our measure of SNR. A typical fMRI experiment has an SNR of approximately one.

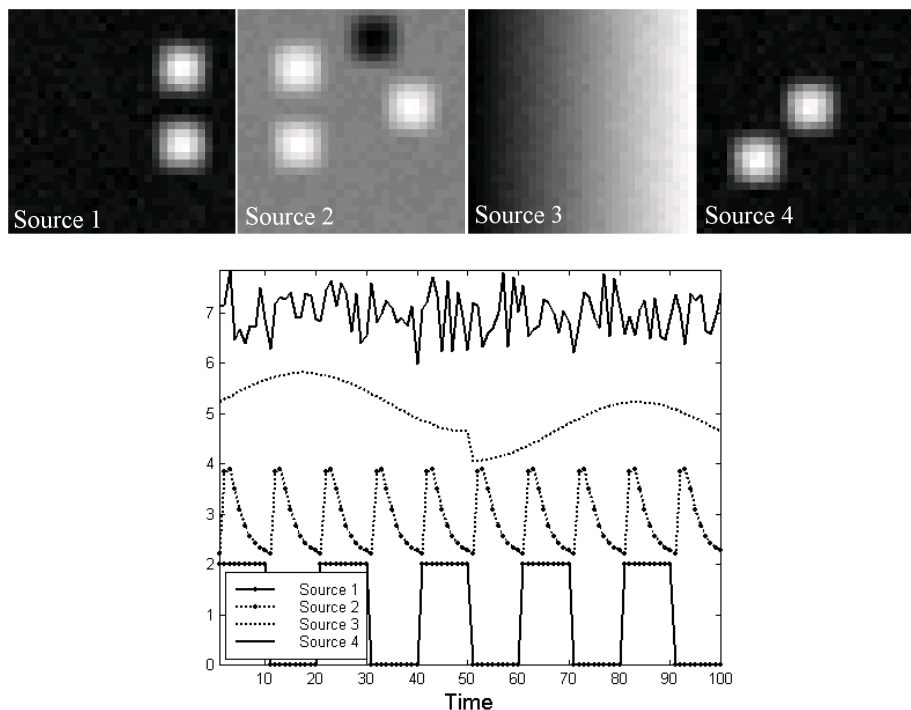


Figure 2: Simulated sources (top), and their mixing functions (bottom). Sources 1 and 2 represent task- and transiently task-related activation, while source 3 represents rotation out of plane resulting in bright voxels on one side and dark voxels on the opposite side. Source 4 simulates function-related activation.

Both ICA algorithms utilize a random starting point for the estimation. We thus performed an ICA estimation scheme twenty times for each of the stages mentioned above in order to capture the variability associated with the starting point as well as the additive noise.

4 RESULTS

The results are depicted in Figure 3 for SNR = 0.75 and SNR = 0.88. The average and standard deviation of the KL divergence we have chosen for evaluation over twenty trials is indicated with black or gray bars, respectively.

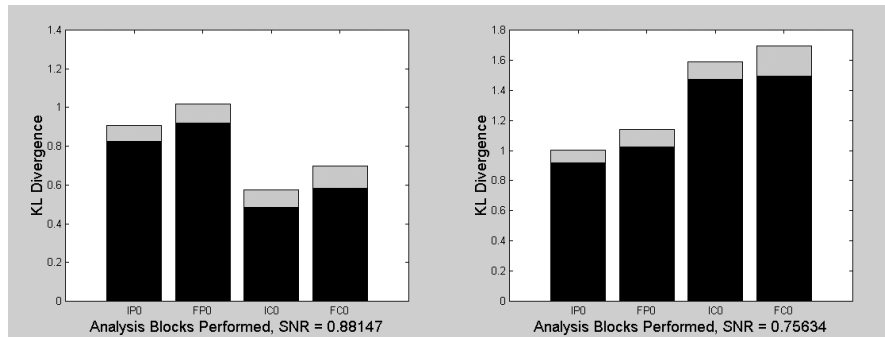


Figure 3: Results from optimization coded by triplets for 1) (F) fastICA or (I) infomax, 2) (P) PCA or (C) clustering, 3) 0/2/4: full width at half maximum smoothing kernel. The mean (std) Kullback-Leibler divergence is displayed in black (gray).

In general, the infomax algorithm performed slightly better than the fastICA algorithm (with the performance difference increasing at higher noise levels) and the clustering performance was better than PCA at lower noise levels. Different types of smoothing did not change the results much and thus only the results for no smoothing are shown. PCA had the least amount of variability in performance while clustering exhibited some variability in performance as it is dependent on the approach used and the initialization. The best combination was infomax with PCA as the data reduction approach.

5 DISCUSSION

The model we present allows addressing performance differences in various stages of the fMRI analysis by ICA. In the data reduction stage for example, a question is how the following stage (ICA) will be impacted by the fact that PCA uses a second order statistics to reduce the dimensionality of the. We have recently shown examples where clustering better preserves the input distribution characteristics

compared to PCA and outperforms PCA [13]. However, our evaluations for this experiment and a number of independent ones pointed in the same direction that clustering performs better for the higher SNR, but at lower SNR the PCA begins to perform better suggesting at increased noise levels the preservation of input distribution characteristics is less critical. It is also true that clustering result is more sensitive to increased noise levels as there is no guarantee for clustering as there is for PCA to obtain the same result each time. However also important to note is the fact that both the high and low SNR experiments were within the range of that for many fMRI experiments and thus it is not clear under which conditions clustering or PCA will perform better for fMRI data. One of our current emphases is the explanation of this tradeoff by comparing various cluster performance measures (e.g. sum-of-squared errors between the original and reduced data, within and between cluster scatter matrix, and higher order criterion such as kurtosis) with similar measures for the PCA decomposition.

The choice of ICA method made less of a difference in the performance. A fundamental difference between the infomax and fastICA algorithms is that the fastICA algorithm forces the output sources to be orthogonal to one another (thus more heavily weighting the second order statistics) whereas infomax does not. Since the higher order statistics are important measures of independence, this may partially explain why infomax performed slightly better than fastICA although closer investigation is needed. Overall smoothing tended to help slightly for all algorithms.

We have presented a synthesis/analysis model for applying independent component analysis to fMRI data. Using this model we generate fMRI-“like” data and optimize the model for this data. We use the Kullback-Leibler divergence between the output sources and the true source distributions as our measure of goodness. Preliminary results suggest that infomax performs slightly better than fastICA and clustering performs better than PCA at low noise levels whereas PCA performs better than clustering at higher noise levels. Both of the noise levels simulated are comparable to what one might encounter in an fMRI experiment. Our work is to our knowledge the first attempt to generate realistic fMRI data and optimize all stages of the ICA processing. It is thus important to perform more simulations and study the theoretical properties of data reduction in more detail, as this choice appears to have the largest effect on the final results.

6 REFERENCES

- [1] K.J.Worsley and K.J.Friston, "Analysis of FMRI Time-Series Revisited-Again," Neuroimage, vol. 2, pp. 173-181, 1995.
- [2] N.Lange, *et al.*, "Plurality and Resemblance in FMRI Data Analysis," Neuroimage, vol. 10, pp. 282-303, 1999.
- [3] G.M.Boynton, *et al.*, "Linear Systems Analysis of Functional Magnetic Resonance Imaging in Human V1," J.Neurosci., vol. 16, pp. 4207-4221, 1996.
- [4] M.J.McKeown, *et al.*, "Analysis of FMRI Data by Blind Separation Into Independent Spatial Components," Hum.Brain Map., vol. 6, pp. 160-188, 1998.
- [5] V.Calhoun, *et al.*, "Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms," Hum.Brain Map., vol. 13, pp. 43-53, 2001.
- [6] A.J.Bell and T.J.Sejnowski, "An Information Maximisation Approach to Blind Separation and Blind Deconvolution," Neural Computation, vol. 7, pp. 1129-1159, 1995. (code at <http://sloan.salk.edu/~tony/ica.html>)
- [7] J.F.Cardoso, "Infomax and Maximum Likelihood for Source Separation," IEEE Letters on Signal Processing, vol. 4, pp. 112-114, 1997.
- [8] A.Hyvarinen and E.Oja, "A Fast Fixed-Point Algorithm for Independent Component Analysis," Neural Computation, vol. 9, pp. 1483-1492, 1997. (code at <http://www.cis.hut.fi/projects/ica/fastica/>)
- [9] D.Malonek, *et al.*, "Vascular Imprints of Neuronal Activity: Relationships Between the Dynamics of Cortical Blood Flow, Oxygenation, and Volume Changes Following Sensory Stimulation," Proc.Natl.Acad.Sci.U.S.A., vol. 94, pp. 14826-14831, 1997.
- [10] M.J.McKeown and T.J.Sejnowski, "Independent Component Analysis of FMRI Data: Examining the Assumptions," Hum.Brain Map., vol. 6, pp. 368-372, 1998.
- [11] R.L.Buckner, "Event-Related FMRI and the Hemodynamic Response," Hum.Brain Map., vol. 6, pp. 373-377, 1998.
- [12] B.Biswal, *et al.*, "Functional Connectivity in the Motor Cortex of Resting Human Brain Using Echo-Planar MRI," Mag.Res.Med., vol. 34, pp. 537-541, 1995.
- [13] V.Calhoun, *et al.*, "Independent Component Analysis Applied To fMRI Data: A Natural Model And Order Selection," Proc. NSIP, Baltimore, 2001.