

Bayesian Analysis of fMRI Data with ICA Based Spatial Prior

Deepti R. Bathula¹, Hemant D. Tagare^{2,3}, Lawrence H. Staib^{2,3}, Xenophon Papademetris^{2,3}, Robert T. Schultz⁴, and James S. Duncan^{1,2,3}

¹ Departments of Biomedical Engineering

² Electrical Engineering

³ Diagnostic Radiology

⁴ Child Study Center

Yale University, P.O. Box 208042, New Haven, CT 06520, USA

deepti.bathula@yale.edu

Abstract. Spatial modeling is essential for fMRI analysis due to relatively high noise in the data. Earlier approaches have been primarily concerned with the spatial coherence of the BOLD response in local neighborhoods. In addition to a smoothness constraint, we propose to incorporate prior knowledge of brain activation patterns learned from training samples. This spatially informed prior can significantly enhance the estimation process by inducing sensitivity to task related regions of the brain. As fMRI data exhibits intersubject variability in functional anatomy, we design the prior using Independent Component Analysis (ICA). Due to the non-Gaussian assumption, ICA does not regress to the mean activation pattern and thus avoids suppressing intersubject differences. Results from a real fMRI experiment indicate that our approach provides statistically significant improvement in estimating activation compared to the standard general linear model (GLM) based methods.

1 Introduction

Functional magnetic resonance imaging (fMRI) enables non-invasive study of brain function by measuring hemodynamic changes caused by neuronal activity. The signal changes observed in fMRI are difficult to measure due to relatively high noise in the data. Consequently, statistical and signal processing techniques are required to reliably detect activation regions. Inspired by segmentation algorithms where prior shape information provides robustness against noise/clutter, we rely on prior knowledge of activation patterns to enhance fMRI analysis.

The general linear model (GLM) [1] is the most commonly used approach for analyzing fMRI data where the time-series of each voxel is statistically compared to an experimental protocol. However, due to low signal-to-noise ratio (SNR), univariate techniques such as GLM tend to produce many small false activation regions. To reduce such false detections, the data are often spatially smoothed using a Gaussian filter in the preprocessing step. Smoothing the data improves

SNR but produces a biased estimate of both the height and the location of activation peaks. Several approaches have been proposed in the literature as alternative denoising techniques to Gaussian smoothing: Markov random fields [2,3], autoregressive spatio-temporal models [4], mixture models [5] and hierarchical Bayesian models with spatial priors [6,7].

While these studies show improvement in sensitivity, the spatial modeling in these techniques is primarily limited to the spatially contiguous and locally homogeneous nature of fMRI responses. In addition to these constraints, the estimation process can be significantly enhanced by incorporating prior knowledge of brain activation patterns (strength, shape & location) that can be learned by analyzing responses to an experiment from a group of subjects. This approach is motivated by scenarios such as pediatric neuroimaging studies of Autism Spectrum Disorders (ASD) [8] where subject motion and discomfort often leads to reduced acquisition periods and noisy data, making it difficult to identify and quantify true activations. Prior knowledge can benefit these subjects by inducing sensitivity to task related regions. The aim of this paper is to design a more spatially informed prior by integrating this knowledge.

A prior for fMRI based on training samples should be constructed carefully, as fMRI data exhibit intersubject variability in brain activation patterns and any prior that does not account for this will invariably cause misleading bias in the estimation. This suggests, for example, that a Principal Component Analysis (PCA) [9] based prior which regresses the posterior toward the mean due to the Gaussian assumption is unsuitable. Further, the prior should be capable of exploiting higher order dependencies in the data - so that the knowledge of coactivating regions may constructively complement the fMRI time-series data.

In this work, we propose to use an Independent Component Analysis (ICA) basis function prior because it has all the required features. Firstly, it can systematically incorporate knowledge of activation patterns learned from training samples into the estimation process. Secondly, it does not exhibit regression to the mean due to non-Gaussian assumption and finally, it is capable of capturing higher order statistical dependencies in the data. We develop a Bayesian framework to incorporate spatial modeling through an ICA based prior and implement temporal modeling using GLM.

2 Description of Method

An fMRI scan produces a set of time series $\{\mathbf{y}_v \in \mathfrak{R}^T : v = 1, \dots, V\}$, each of which represents the measured BOLD signal over T time samples in one of V voxels. The General Linear Model (GLM) assumes that the BOLD signal is a linear combination of protocol-dependent components and noise. Assuming Gaussian white noise, the voxel-wise GLM is written as

$$\mathbf{y}_v = \mathbf{X}\beta_v + \epsilon_v \quad \epsilon_v \sim \mathcal{N}(0, \lambda_v^{-1}I_T) \quad (1)$$

where ϵ_v is the residual, λ_v denotes the noise precision and β_v are regression coefficients on the columns of the design matrix \mathbf{X} containing R regressors.

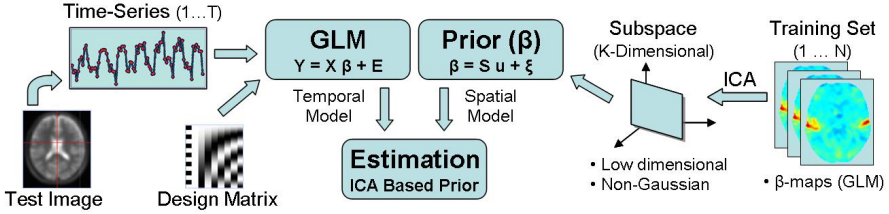


Fig. 1. A schematic of the proposed estimation method using ICA based prior

Collecting y_v and β_v over all voxels, the above equation can be written succinctly in matrix notation as $\mathbf{Y} = \mathbf{X}\beta + \mathbf{E}$.

The assumption that noise follows an i.i.d. Gaussian distribution is only an approximation due to the presence of temporal autocorrelations, but in this work we focus on spatial modeling. However, our model can be easily extended to deal with serial correlation using autoregressive processes as in [6].

2.1 Bayesian Formulation

The fundamental idea of this work is to exploit prior knowledge of brain activation patterns to complement the information provided by the functional time-series. This *a priori* information can be naturally incorporated into a Bayesian framework by specifying spatial priors over the regression coefficients. Temporal modeling of the data is implemented using GLM as described below. A graphical representation of the proposed estimation model is given in Figure 1.

Likelihood Model

Using the same model for the BOLD signal as in Eq. (1), the conditional probability of obtaining an fMRI scan, \mathbf{Y} , given the activation strength of all voxels, β , can be expressed as:

$$p(\mathbf{Y}|\beta, \lambda) = \prod_{v=1}^V p(y_v|\beta_v, \lambda_v) = \prod_{v=1}^V \mathcal{N}(y_v; \mathbf{X}\beta_v, \lambda_v^{-1} I_T) \tag{2}$$

where the equalities are obtained based on the assumption that, given the activation strength vector of each voxel β_v , the fMRI time-series of different voxels are conditionally independent. We model spatial coherence into activation parameters and not the observed time-series.

Prior Model

Prior information learned from training data is increasingly used in image analysis and computer vision. For a specific fMRI experiment, analyzing responses evoked by a group of subjects can provide valuable information about brain regions relevant to the task as well as the magnitude, shape and location of activations. As the signal change due to the BOLD effect is very small and noisy,

prior knowledge of activation patterns can significantly aid the estimation process. However, due to the presence of intersubject variability in brain activation patterns, designing such a prior for fMRI data needs careful consideration.

In this work, we prefer Independent Component Analysis (ICA) for the following properties. As ICA produces spatially localized and statistically independent basis vectors, variations in the shape and location of activations can be captured effectively. Unlike Principal Component Analysis (PCA), which assumes a Gaussian distribution of patterns [9], ICA does not regress to the mean activation pattern and thus avoids suppressing intersubject differences in functional anatomy. In addition to covariance between voxels, ICA is also sensitive to higher-order dependencies in the data. Hence, activations retained within each component are statistically dependent and this knowledge can constructively complement the functional time-series data by avoiding excessive blurring of small activation regions that coactivate with other regions.

As learning from training data requires accurate alignment of subjects, individual brains are nonrigidly transformed into Talairach space. Given a set of coregistered and realigned regression coefficient (or activation strength) maps, ICA generates patterns that are maximally statistically independent. Generally components reflecting noise are eliminated, retaining only the relevant components. For each regressor r , we represent the K retained source patterns by columns of matrix S_r (indicating the non-Gaussian subspace). Any regression coefficient image β_r can then be represented as a linear combination of these components, $\beta_r = S_r u_r + \xi_r$, where u_r is a vector of mixing coefficients and ξ_r is the error term. Assuming i.i.d. Gaussian distribution for residuals ξ_r , the spatial prior over regression coefficients is given by

$$p(\beta|\mathbf{U}, \boldsymbol{\tau}) = \prod_{r=1}^R p(\beta_r|u_r, \tau_r) = \prod_{r=1}^R \mathcal{N}(\beta_r; \mathbf{S}_r u_r, \tau_r^{-1} I_N) \tag{3}$$

where τ_r is the precision of the residuals for regressor r .

Mixing Coefficients - The unknown mixing coefficients \mathbf{U} are assumed to be independent and identically distributed (i.i.d) and a prior on them is defined using an exponential power (EP) distribution.

$$p(\mathbf{U}) = \prod_{r=1}^R \prod_{k=1}^K p(u_{r,k}), \quad p(u_{r,k}) = EP(b_0, c_0) \tag{4}$$

As the ICA basis functions are ambiguous in sign and scaling, the shape parameter c_0 of the density function is set to 0.5 such that the distribution is approximately uniform in the empirically chosen range $(-b_0, +b_0)$.

Precisions - As in [7], we use Gamma priors on the observation noise precisions $\boldsymbol{\lambda}$ (Eq. 2) and ICA residuals $\boldsymbol{\tau}$ (Eq. 3)

$$p(\boldsymbol{\lambda}) = \prod_{v=1}^V p(\lambda_v) = \prod_{v=1}^V Ga(\lambda_v; b_1, c_1), \quad p(\boldsymbol{\tau}) = \prod_{r=1}^R p(\tau_r) = \prod_{r=1}^R Ga(\tau_r; b_2, c_2) \tag{5}$$

Gamma priors are chosen as they are the standard conjugate priors for Gaussian error models. The values of hyperparameters b_1, c_1, b_2, c_2 are set so as to specify vague or uninformative priors due to lack of knowledge about $\boldsymbol{\lambda}$ and $\boldsymbol{\tau}$.

Maximum a Posteriori Estimate

The posterior distribution of unknown parameters $\Theta = \{\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{U}, \boldsymbol{\tau}\}$ given the fMRI time-series data \mathbf{Y} can be written using Bayes' rule as

$$p(\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{U}, \boldsymbol{\tau} | \mathbf{Y}) \propto p(\mathbf{Y} | \boldsymbol{\beta}, \boldsymbol{\lambda}) p(\boldsymbol{\lambda}) p(\boldsymbol{\beta} | \mathbf{U}, \boldsymbol{\tau}) p(\boldsymbol{\tau}) p(\mathbf{U}) \quad (6)$$

or more concisely $p(\Theta | \mathbf{Y}) \propto p(\mathbf{Y} | \Theta) p(\Theta)$. Taking the natural logarithm, which is a monotonically increasing function, the maximum *a posteriori* (MAP) estimate of Θ can be expressed as

$$\Theta_{map} = \arg \max_{\Theta} [\ln p(\mathbf{Y} | \Theta) + \alpha \ln p(\Theta)] = \arg \min_{\Theta} [E(\Theta)] \quad (7)$$

where $E(\Theta) = -[\ln p(\mathbf{Y} | \Theta) + \alpha \ln p(\Theta)]$ is the corresponding posterior energy function and α is the scaling factor for the prior term. A closed form solution for equation (7) does not exist. Therefore, we minimize $E(\Theta)$ using coordinate descent which iteratively minimizes the energy function for each unknown parameter until convergence.

3 Experimental Results

We applied the proposed estimation method to an attention modulation experiment conducted on 12 healthy subjects. Imaging was performed with a whole-body Siemens Trio 3T scanner. Functional T2*-weighted images were acquired using a gradient-echo EPI sequence (40 axial slices, 64x64 matrix, voxel size = 3.5 x 3.5 x 3.5 mm³ with no gap, TR/TE = 2320/25 ms). The experimental paradigm consisted of an ON-OFF block design where the subject was shown images of either houses or faces during the ON phase and asked to view a fixation point during the OFF phase. The experiment consisted of 5 runs with each run generating 140 time samples. Data preprocessing involved motion correction and linear detrending but *no* spatial smoothing.

Validation of analysis techniques for real fMRI data is difficult due to the lack of ground truth. However, to evaluate the performance of different estimation methods we consider the activation strength estimated from the entire time-series (all 5 runs) using GLM as ground truth or reference and compare it to the estimates from reduced time-series (2 runs).

Our estimation method with ICA based prior is tested using the leave-one-out technique where the prior model is trained over 11 subjects and tested on the 12th subject. We use the Matlab implementation of FastICA [10] for performing ICA and empirically choose the components to be retained as there was a clear distinction between relevant and noisy components. For large datasets, measures of non-Gaussianity can be used to automatically select relevant components.

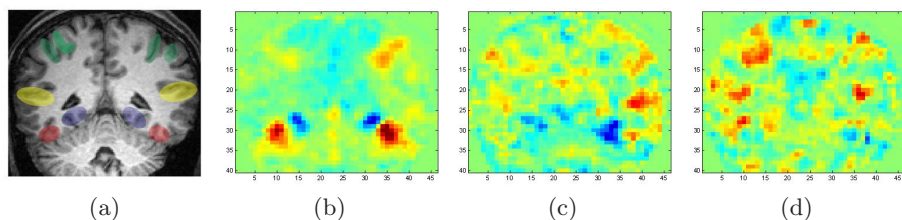


Fig. 2. (a) Structural image of a coronal slice with task related brain regions highlighted: Blue - Parahippocampal Place Areas (PPA), Red - Fusiform Face Areas (FFA), Yellow - Superior Temporal Sulcus (STS) and Green - Intraparietal Sulcus (IPS); (b),(c),(d) - ICA basis vectors describing independent activation patterns

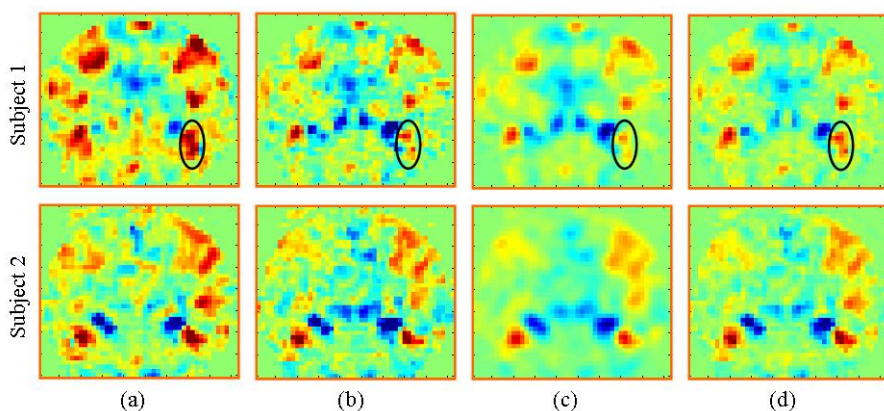


Fig. 3. Activation strength maps for the main effect of face versus house in a coronal slice of two subjects (top & bottom rows). Images of estimated contrast using: (a) GLM on 5-run data (reference) (b) GLM on 2-run data (c) GLM on 2-Run data smoothed by a 6mm FWHM Gaussian kernel (d) ICA based prior on 2-Run data ($K = 4, \alpha = 0.8$).

Figure 2 displays three of the independent components (ICs) produced from this data set for a coronal slice. The first component (Fig. 2b) captures the standard activation for face versus house contrast with positive activation in Fusiform Face Areas (FFA) and negative activation in Parahippocampal Place Areas (PPA). The remaining components (Fig. 2c & 2d) show activation in other face-sensitive areas as observed in some of the subjects. These task related areas are highlighted on the structural scan in Figure 2(a).

In this work, we compare the estimate of activation strength using ICA based prior with GLM estimate and smoothed-GLM estimate where the data is smoothed with a modest 6mm FWHM Gaussian kernel. For a visual comparison, estimated maps of face versus house contrast in a coronal slice from two subjects are shown in Figure 3. Considering Subject 1 (top row), the reference map (Fig. 3a) shows a typical activation for this task in PPA and FFA along with positive activation in Superior Temporal Sulcus (STS) and Intraparietal

Table 1. Effectiveness of different estimators measured using Sum-of-Squares Error (SSE) and Correlation Coefficient (ρ) between estimated activation strength maps and reference maps for $N = 12$ normal subjects

	GLM	Smoothed-GLM	ICA Based Prior
SSE	40 ± 19	37 ± 16	24 ± 10
ρ	0.69 ± 0.11	0.68 ± 0.11	0.80 ± 0.09

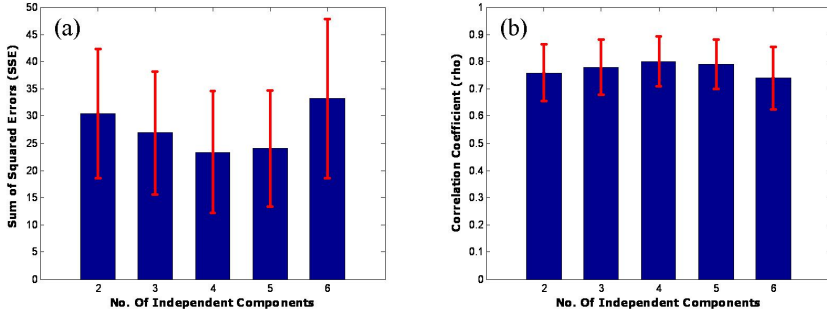


Fig. 4. Effect of changing the number of retained ICs measured using Sum-of-Squares Error (SSE) and Correlation Coefficient (ρ) between estimated activation strength maps and reference maps for $N = 12$ normal subjects

Sulcus (IPS). The contrast map obtained from GLM on 2-run data (Fig. 3b) also shows a similar activation pattern but lower in magnitude and extent. It is also very noisy and fragmented, especially in the right FFA (highlighted). Smoothed-GLM on 2-run data (Fig. 3c) considerably reduced background noise but also lowered the activation strength in task-sensitive areas and almost eliminated activation in the right FFA. The effect of introducing the ICA based prior, with number of retained components $K = 4$ and scaling factor $\alpha = 0.8$, into the estimation process is shown in Figure 3(d). It provides reasonable smoothing in the background while avoiding excessive blurring in task-sensitive areas, which is achieved by retaining only the relevant basis vectors. And unlike smoothed-GLM, prior knowledge of activation patterns enables it to recover activation in the right FFA which is also observed in the reference map as highlighted in Figure 3(a). The bottom row of Figure 3 displays the estimated contrast maps for a different subject. Again, ICA based prior provides significant improvement in estimated activation by reducing background noise and accurately emphasizing activation in task-sensitive regions.

A quantitative comparison of these estimators is provided in Table 1. The effectiveness of the estimators is measured using sum-of-squares error (SSE) and correlation coefficient (ρ) between the estimated maps and the reference map. Table 1 shows the mean and standard deviation of these quantities obtained by analyzing the same coronal slice (Fig. 2a) of all 12 subjects using leave-one-out technique. The GLM estimate has the most background noise and also has the

highest SSE and lowest ρ . As expected, spatial smoothing with a fixed Gaussian kernel decreases SSE but also reduces ρ . Using the ICA based prior reduced the mean SSE by about 40% and increased ρ by 0.11 relative to the GLM estimate. The SSE results indicate that ICA based prior provides statistically significant ($p < 0.001$, paired t-test) improvement over GLM and smoothed-GLM.

Figure 4 displays the sensitivity of the ICA based estimation method to the number of retained ICs (K). An underestimate of the number of components generally leaves some of the patterns unrecognized, whereas an overestimate leads to an increased fraction of components reflecting noise. This trend is exhibited in the graphs which show the optimal number of components to be four with relatively stable performance in the range of 3 to 5 components. Finally, an experiment with varying scale factor (α) indicated that the performance of the estimation model was fairly robust to the changes in α value.

4 Discussion

In this paper, we presented a Bayesian approach to spatial modeling of fMRI data that embodies prior knowledge of brain activation patterns learned from training samples. The ICA based prior not only captures activation patterns effectively, but also accommodates intersubject variability due to the non-Gaussian assumption. A comparison with the standard GLM based methods for real fMRI data demonstrates that the algorithm provides statistically significant improvement in estimating activation. We plan to apply the proposed method to an Autism study where subject discomfort often leads to reduced acquisition periods and noisy data, which can significantly benefit from a spatially informed prior that is sensitive to task related regions. Future efforts will be directed towards integrating temporal correlations into the Bayesian framework and exploiting anatomical information for a more robust, sensitive and accurate estimation model.

References

1. Friston, K.J., Holmes, A.P.: Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping* 2(4), 189–210 (1994)
2. Descombes, X., Kruggel, F., von Cramon, Y.: fMRI signal restoration using an edge preserving spatio-temporal Markov random field. *NeuroImage* 8, 340–349 (1998)
3. Rajapakse, J.C., Piyaratna, J.: Bayesian approach to segmentation of statistical parametric maps. *IEEE TMI* 48(10), 1186–1194 (2001)
4. Woolrich, M.W., Jenkinson, M., Brady, J.M., Smith, S.M.: Fully Bayesian spatio-temporal modeling of fMRI data. *IEEE TMI* 23(2), 213–231 (2004)
5. Hartvig, N., Jensen, J.: Spatial mixture modeling of fMRI data. *Human Brain Mapping* 11(4), 233–248 (2000)
6. Penny, W.D., Trujillo-Barreto, N.J., Friston, K.J.: Bayesian fMRI time series analysis with spatial priors. *NeuroImage* 24, 350–362 (2005)
7. Flandin, G., Bayesian, W.D.P.: fMRI data analysis with sparse spatial basis function priors. *NeuroImage* 34, 1108–1125 (2006)

8. Schultz, R.T.: Developmental deficits in social perception in autism: The role of the amygdala and fusiform face area. *Intl. J. Dev. Neuroscience* 23, 125–141 (2005)
9. Yang, J., Papademetris, X., Staib, L.H., Duncan, J.S.: Functional brain image analysis using joint function-structure priors. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) *MICCAI 2004. LNCS*, vol. 3217, pp. 736–744. Springer, Heidelberg (2004)
10. Hyvriinen, A., Oja, E.: A fast fixed-point algorithm for independent component analysis. *Neural Computation* 9(7), 1483–1492 (1997)