Chapter 25: Spatio-temporal models for fMRI

W. Penny, G. Flandin and N. Trujillo-Barreto

May 8, 2006

Introduction

Functional Magnetic Resonance Imaging (fMRI) using Blood Oxygen Level Dependent (BOLD) contrast is an established method for making inferences about regionally specific activations in the human brain [Frackowiak et al. 2003]. From measurements of changes in blood oxygenation one can use various statistical models, such as the General Linear Model (GLM) [Friston et al. 1995], to make inferences about task-specific changes in underlying neuronal activity.

This chapter reviews previous work [Penny et al. 2003, Penny and Flandin 2005, Penny et al. 2005c, Penny et al. 2006] on the development of Spatially Regularised General Linear Models (SRGLMs) for the analysis of fMRI data. These models allow for the characterisation of subject and regionally specific effects using spatially regularised Posterior Probability Maps (PPMs). This spatial regularisation has been shown [Penny et al. 2005c] to increase the sensitivity of inferences one can make.

The chapter is structured as follows. The theoretical section describes the generative model for SRGLM. This is split into descriptions of the prior and likelihood. We show how the Variational Bayes (VB) algorithm, described in the previous chapter, can be used for approximate inference. We describe how these inferences are implemented for uni- and multi-variate contrasts, and discuss the rationale for thresholding the resulting PPMs. We also discuss the spatio-temporal nature of the model and compare it with standard approaches. The results section looks at null fMRI data, synthetic data and fMRI from functional activation studies of auditory and face processing. The chapter finishes with a discussion.

Notation

Lower case variable names denote vectors and scalars. Whether the variable is a vector or scalar should be clear from the context. Upper case names denote matrices or dimensions of matrices. In what follows $\mathsf{N}(x;\mu,\Sigma)$ denotes a multivariate normal density over x, having mean μ and covariance Σ. The precision of a Gaussian variate is the inverse (co)variance. A gamma density over the scalar random variable x is written as $\mathsf{Ga}(x;a,b)$. Normal and Gamma densities are defined in chapter 24. We also use $||x||^2 = x^T x$, denote the trace operator as $\mathsf{Tr}(X)$, X^+ for the pseudo-inverse, and use $\mathsf{diag}(x)$ to denote a diagonal matrix with diagonal entries given by the vector x.

Theory

We denote an fMRI data set consisting of T time points at N voxels as the $T \times N$ matrix Y. In mass-univariate models [Friston et al. 1995], these data are explained in terms of a $T \times K$ design matrix X, containing the values of K regressors at T time points, and a $K \times N$ matrix of regression coefficients W, containing K regression coefficients at each of the N voxels. The model is written

$$Y = XW + E \tag{1}$$

where E is a $T \times N$ error matrix.

It is well known that fMRI data is contaminated with artifacts. These stem primarily from low-frequency drifts due to hardware instabilities, aliased cardiac pulsation and respiratory sources, unmodelled neuronal activity and residual motion artifacts not accounted for by rigid body registration methods [Woolrich et al. 2001]. This results in the residuals of an fMRI analysis being temporally autocorrelated.

In previous work we have shown that, after removal of low-frequency drifts using Discrete Cosine Transform (DCT) basis sets, low-order voxel-wise autoregressive (AR) models are sufficient for modelling this autocorrelation [Penny et al. 2003]. It is important to model these noise processes because parameter estimation then becomes less biased [Gautama and Van Hulle 2004] and more accurate [Penny et al. 2003].

Model likelihood

We now describe the approach taken in our previous work. For a Pth-order AR model, the likelihood of the data is given by

$$p(Y|W, A, \lambda) = \prod_{t=P+1}^{T} \prod_{n=1}^{N} \mathsf{N}(y_{tn} - x_t w_n; (d_{tn} - X_t w_n)^T a_n, \lambda_n^{-1}) \tag{2}$$

where n indexes the nth voxel, a_n is a $P \times 1$ vector of autoregressive coefficients, w_n is a $K \times 1$ vector of regression coefficients and λ_n is the observation noise precision. The vector x_t is the tth row of the design matrix and X_t is a $P \times K$ matrix containing the previous P rows of X prior to time point t. The scalar y_{tn} is the fMRI scan at the tth time point and nth voxel and $d_{tn} = [y_{t-1,n}, y_{t-2,n}, ..., y_{t-P,n}]^T$. Because d_{tn} depends on data P time steps before, the likelihood is evaluated starting at time point $P + 1$, thus ignoring the GLM fit at the first P time points.

Equation 2 shows that higher model likelihoods are obtained when the prediction error $y_{tn} - x_t w_n$ is closer to what is expected from the AR estimate of prediction error.

The voxel-wise parameters w_n and a_n are contained in the nth columns of matrices W and A, and the voxel-wise precision λ_n is the nth entry in λ. The next section describes the prior distributions over these parameters. Together, the likelihood and prior define the generative model, which is shown in Figure 1.

Priors

The graph in Figure 1 shows that the joint probability of parameters and data can be written

$$\begin{aligned} p(Y,W,A,\lambda,\alpha,\beta) &= p(Y|W,A,\lambda)p(W|\alpha) \\ &\quad p(A|\beta)p(\lambda|u_1,u_2) \\ &\quad p(\alpha|q_1,q_2)p(\beta|r_1,r_2) \end{aligned} \tag{3}$$

where the first term is the likelihood and the other terms are the priors. The likelihood is given in equation 2 and the priors are described below.

Regression coefficients

For the regressions coefficients we have

$$\begin{aligned} p(W|\alpha) &= \prod_{k=1}^{K} p(w_k^T|\alpha_k) \\ p(w_k^T|\alpha_k) &= N(w_k^T;0,\alpha_k^{-1}D_w^{-1}) \end{aligned} \tag{4}$$

where D_w is a spatial precision matrix. This can be set to correspond to eg. a Low Resolution Tomography (LORETA) prior, a Gaussian Markov Random Field (GMRF) prior or a minimum norm prior ($D_w = I$) [Friston and Penny 2003] as described in earlier work [Penny et al. 2005c]. These priors implement the spatial regularisation and are specified separately for each slice of data. Specification of 3-dimensional spatial priors (ie. over multiple slices) is desirable from a modelling perspective but is computationally too demanding for current computer technology.

We can also write $w_v = vec(W)$, $w_r = \mathsf{vec}(W^T)$, $w_v = H_w w_r$ where H_w is a permutation matrix. This leads to

$$\begin{aligned} p(W|\alpha) &= p(w_v|\alpha) \\ &= N(w_v;0,B^{-1}) \end{aligned} \tag{5}$$

where B is an augmented spatial precision matrix given by

$$B = H_w(\mathsf{diag}(\alpha) \otimes D_w)H_w^T \tag{6}$$

where $\otimes$ is the Kronecker product. This form of the prior is useful as our specification of approximate posteriors is based on similar quantities. It can be seen that α encodes the spatial precision of the regression coefficients.

The above Gaussian priors underly GMRFs and LORETA and have been used previously in fMRI [Woolrich et al. 2004] and EEG [Pascal Marqui et al. 1994]. They are by no means, however, the optimal choice for imaging data. In EEG, for example, much interest has focussed on the use of L^p-norm priors [Auranen 2005] instead of the L^2-norm implicit in the Gaussian assumption. Additionally, we are currently investigating the use of wavelet priors. This is an active area of research and will be the topic of future publications.

AR coefficients

We also define a spatial prior for the AR coefficients so that they too can be spatially regularised. We have

$$\begin{aligned} p(A|\beta) &= \prod_{p=1}^{P} p(a_p|\beta_p) \\ p(a_p|\beta_p) &= N(a_p; 0, \beta_p^{-1} D_a^{-1}) \end{aligned} \tag{7}$$

Again, D_a is a user-defined spatial precision matrix, $a_v = \mathsf{vec}(A)$, $a_r = \mathsf{vec}(A^T)$ and $a_v = H_a a_r$ where H_a is a permutation matrix. This prior is used to implement the spatial regularisation of the AR coefficients. We can write

$$\begin{aligned} p(A|\beta) &= p(a_v|\beta) \\ &= N(a_v; 0, J^{-1}) \end{aligned} \tag{8}$$

where J is an augmented spatial precision matrix

$$J = H_a \left(\mathsf{diag}(\beta) \otimes D_a\right) H_a^T \tag{9}$$

This form of the prior is useful as our specification of approximate posteriors is based on similar quantities. The parameter β plays a similar role as α and controls the spatial regularisation of the temporal auto-regression coefficients.

We have also investigated 'Tissue-type' priors which constrain AR estimates to be similar for voxels in the same tissue-type eg. gray matter, white matter or cerebro-spinal fluid. Bayesian model selection [Penny et al. 2006], however, favours the smoothly varying priors defined in equation 7.

Precisions

We use Gamma priors on the precisions α, β and λ

$$\begin{aligned} p(\lambda|u_1, u_2) &= \prod_{n=1}^{N} \mathsf{Ga}(\lambda_n; u_1, u_2) \\ p(\alpha|q_1, q_2) &= \prod_{k=1}^{K} \mathsf{Ga}(\alpha_k; q_1, q_2) \\ p(\beta|r_1, r_2) &= \prod_{p=1}^{P} \mathsf{Ga}(\beta_p; r_1, r_2) \end{aligned} \tag{10}$$

where the Gamma density is defined in Chapter 24. Gamma priors were chosen as they are the conjugate priors for Gaussian error models. The parameters are set to $q_1 = r_1 = u_1 = 10$ and $q_2 = r_2 = u_2 = 0.1$. These parameters produce Gamma densities with a mean of 1 and a variance of 10. The robustness of, for example, model selection to the choice of these parameters is discussed in [Penny et al. 2003].

Approximate Posteriors

Inference for SRGLMs has been implemented using the Variational Bayes (VB) approach described in the previous chapter. In this section we describe the

algorithm developed in previous work [Penny et al. 2005c] where we assumed that the approximate posterior factorises over voxels and subsets of parameters.

Because of the spatial priors, the regression coefficients in the true posterior $p(W|Y)$ will clearly be correlated. Our perspective, however, is that this is too computationally burdensome for current personal computers to take account of. Moreover, as we shall see later, updates for the approximate factorised densities $q(w_n)$ do encourage the approximate posterior means to be similar at nearby voxels, thereby achieving the desired effect of the prior.

Our approximate posterior is given by

$$\begin{aligned} q(W, A, \lambda, \alpha, \beta) &= \prod_n q(w_n) q(a_n) q(\lambda_n) \\ &\quad \prod_k q(\alpha_k) \prod_p q(\beta_p) \end{aligned} \tag{11}$$

and each component of the approximate posterior is described below. These update equations are self-contained except for a number of quantities that are marked out using the 'tilde' notation. These are $\tilde{A}_n, \tilde{b}_n, \tilde{C}_n, \tilde{d}_n$ and $\tilde{G}_n$ which are all defined in the Appendix of this Chapter.

Regression coefficients

We have

$$\begin{aligned} q(w_n) &= \mathsf{N}(w_n; \hat{w}_n, \hat{\Sigma}_n) \\ \hat{\Sigma}_n &= \left(\bar{\lambda}_n \tilde{A}_n + \bar{B}_{nn} \right)^{-1} \\ \hat{w}_n &= \hat{\Sigma}_n \left(\bar{\lambda}_n \tilde{b}_n^T + r_n \right) \\ r_n &= - \sum_{i=1, i \neq n}^{N} \bar{B}_{ni} \hat{w}_i \end{aligned} \tag{12}$$

where $\hat{w}_n$ is the estimated posterior mean and $\hat{\Sigma}_n$ is the estimated posterior covariance. The quantity $\bar{B}$ is defined as in equation 6 but uses $\bar{\alpha}$ instead of α. The quantities $\tilde{A}_n$ and $\tilde{b}_n$ are expectations related to autoregressive processes and are defined in the Appendix. In the absence of temporal autocorrelation we have $\tilde{A}_n = X^T X$ and $\tilde{b}_n^T = X^T y_n$.

AR coefficients

We have

$$q(a_n) = \mathsf{N}(a_n; m_n, V_n)$$

where

$$\begin{aligned} V_n &= \left(\bar{\lambda}_n \tilde{C}_n + \bar{J}_{nn} \right)^{-1} \\ m_n &= V_n (\bar{\lambda}_n \tilde{d}_n + j_n) \\ j_n &= - \sum_{i=1, i \neq n}^{N} \bar{J}_{ni} m_i \end{aligned} \tag{13}$$

and m_n is the estimated posterior mean and V_n is the estimated posterior covariance. The quantity $\bar{J}$ is defined as in equation 9 but $\bar{\beta}$ is used instead of β. The subscripts in $\bar{J}_{ni}$ denote that part of $\bar{J}$ relevant to the nth and ith voxels. The quantities $\tilde{C}_n$ and $\tilde{d}_n$ are expectations that are defined in the Appendix.

Precisions

The approximate posteriors over the precision variables are Gamma densities. For the precisions on the observation noise we have

$$\begin{aligned} q(\lambda_n) &= \mathsf{Ga}(\lambda_n; b_n, c_n) \\ \frac{1}{b_n} &= \frac{\tilde{G}_n}{2} + \frac{1}{u_1} \\ c_n &= \frac{T}{2} + u_2 \\ \bar{\lambda}_n &= bc \end{aligned} \tag{14}$$

where $\tilde{G}_n$ is the expected prediction error defined in the Appendix. For the precisions of the regression coefficients we have

$$\begin{aligned} q(\alpha_k) &= \mathsf{Ga}(\alpha_k; g_k, h_k) \\ \frac{1}{g_k} &= \frac{1}{2}\left(\mathsf{Tr}(\hat{\Sigma}_k D_w) + \hat{w}_k^T D_w \hat{w}_k\right) + \frac{1}{q_1} \\ h_k &= \frac{N}{2} + q_2 \\ \bar{\alpha}_k &= g_k h_k \end{aligned} \tag{15}$$

For the precisions of the AR coefficients we have

$$\begin{aligned} q(\beta_p) &= \mathsf{Ga}(\beta_p; r_{1p}, r_{2p}) \\ \frac{1}{r_{1p}} &= \frac{1}{2}\left(\mathsf{Tr}(V_p D_a) + m_p^T D_a m_p\right) + \frac{1}{r_1} \\ r_{2p} &= \frac{N}{2} + r_2 \\ \bar{\beta}_p &= r_{1p} r_{2p} \end{aligned} \tag{16}$$

Practicalities

Our empirical applications use spatial precision matrices D_a and D_w, defined in section 2.2, which produce GMRF priors. Also, we use AR models of order $P = 3$. Model selection using VB showed that this model order was sufficient for all voxels in a previous analysis of fMRI [Penny et al. 2003].

The VB algorithm is initialised using Ordinary Least Square (OLS) estimates for regression and autoregressive parameters as described in [Penny et al. 2003]. Quantities are then updated using equations 12,13,14,15,16.

As described in the previous chapter, the aim of VB is to match an approximate posterior to the true posterior density in the sense of minimising Kullback-Liebler (KL) divergence. This is implemented implicitly by maximising the quantity F, known in statistical physics as the negative free energy.

In the implemention of VB for SRGLMs, F is monitored during optimisation. Convergence is then defined as less than a 1% increase in F.

Expressions for computing F are given in [Penny et al. 2006]. This is an important quantity as it can also be used for model comparison. This is described at length in [Penny et al. 2006] and reviewed in Chapters 24 and 35.

The algorithm we have described is implemented in SPM version 5 and can be downloaded from [SPM Software 2006]. Computation of a number of quantites (eg. $\tilde{C}_n$, $\tilde{d}_n$ and $\tilde{G}_n$) is now much more efficient than in previous versions [Penny et al. 2005c]. These improvements are described in a separate document [Penny and Flandin 2005b]. To analyse a single session of data (eg. the face fMRI data) takes about 30 minutes on a typical personal computer.

Spatio-temporal deconvolution

The central quantity of interest in fMRI analysis is our estimate of effect sizes, embodied in contrasts of regression coefficients. A key update equation in our VB scheme is, therefore, the approximate posterior for the regression coefficients. This is given by equation 12. For the special case of temporally uncorrelated data we have

$$\begin{aligned} \hat{\Sigma}_n &= \left(\bar{\lambda}_n X^T X + \bar{B}_{nn}\right)^{-1} \\ \hat{w}_n &= \hat{\Sigma}_n \left(\bar{\lambda}_n X^T y_n + r_n\right) \end{aligned} \tag{17}$$

where $\bar{B}$ is a spatial precision matrix and r_n is the weighted sum of neighboring regression coefficient estimates.

This update indicates that the estimate at a given voxel regresses towards those at nearby voxels. This is the desired effect of the spatial prior and is preserved despite the factoristion over voxels in the approximate posterior (see equation 11). Equation 17 can be thought of as the combination of a temporal prediction $X^T y_n$ and a spatial prediction from r_n. Each prediction is weighted by its relative precision to produce the optimal estimate $\hat{w}_n$. In this sense, VB provides a spatio-temporal deconvolution of fMRI data. Moreover, the parameters controlling the relative precisions, $\bar{\lambda}_n$ and $\bar{\alpha}$ are estimated from the data. The effect size estimates therefore derive from an automatically regularised spatio-temporal deconvolution.

Contrasts

After having estimated a model, we will be interested in characterising a particular effect, c, which can usually be expressed as a linear function or 'contrast' of parameters, w. This is described at length in chapter 9. That is,

$$c_n = C^T w_n \tag{18}$$

where C is a contrast vector or matrix. For example, the contrast vector $C^T = [1, -1]$ computes the difference between two experimental conditions.

Our statistical inferences are based on the approximate distribution $q(W)$, which implies a distribution on c, $q(c)$. Because $q(W)$ factorises over voxels we can write

$$q(c) = \prod_{n=1}^{N} q(c_n) \tag{19}$$

where c_n is the effect size at voxel n. Given a contrast matrix C we have

$$q(c_n) = \mathsf{N}(c_n; \mu_n, S_n) \tag{20}$$

with mean and covariance

$$\begin{aligned} \mu_n &= C^T \hat{w}_n \\ S_n &= C^T \hat{\Sigma}_n C \end{aligned} \tag{21}$$

Bayesian inference based on this posterior can then take place using confidence intervals [Box and Tiao 1992]. For univariate contrasts we have suggested the use of Posterior Probability Maps (PPMs), as described in chapter 23.

If c_n is a vector then we have a multivariate contrast. Inference can then proceed as follows. The probability α that the zero vector lies on the $1-\alpha$ confidence region of the posterior distribution at each voxel must then be computed. We first note that this probability is the same as the probability that the vector μ_n lies on the edge of the $1-\alpha$ confidence region for the distribution $N(\mu_n; 0, S_n)$. This latter probability can be computed by forming the test statistic

$$d_n = \mu_n^T S_n^{-1} \mu_n \tag{22}$$

which will be the sum of $r_n = rank(S_n)$ independent, squared Gaussian variables. As such, it has a χ^2 distribution

$$p(d_n) = \chi^2(r_n) \tag{23}$$

This procedure is identical to that used for making inferences in Bayesian Multivariate Autoregressive Models [Harrison et al. 2003]. We can also use this procedure to test for two-sided effects, that is, activations or deactivations. Though, strictly, these contrasts are univariate we will use the term 'multivariate contrasts' to cover these two-sided effects.

Thresholding

In previous work [Friston and Penny 2003] we have suggested deriving Posterior Probability Maps (PPMs) by applying two thresholds to the posterior distributions (i) an effect size threshold, γ, and (ii) a probability threshold p_T. Voxel n is then included in the PPM if $q(c_n > \gamma) > p_T$. This approach was described in Chapter 23.

If voxel n is to be included, then the posterior exceedance probability $q(c_n > \gamma)$ is plotted. It is also possible to plot the effect size itself, c_n. The following exploratory procedure can be used for exploring the posterior distribution of effect sizes. Firstly, plot a map of effect sizes using the thresholds $\gamma = 0$ and $p_T = 1 - 1/N$ where N is the number of voxels. We refer to these values as the 'default thresholds'. Then, after visual inspection of the resulting map use a non-zero γ, the value of which reflects effect sizes in areas of interest. It will then be possible to reduce p_T to a value such as 0.95. Of course, if previous imaging analyses have indicated what effect sizes are physiologically relevant then this exploratory procedure is unnecessary.

False positive rates

If we partition effect-size values into two hypothesis spaces $H_0 : c \leq \gamma$ and $H_1 : c > \gamma$ then we can characterise the sensitivity and specificity of our algorithm. This is different to classical inference which uses $H_0 : c = 0$. A False Positive (FP) occurs if we accept H_1 when H_0 is true.

If we use the default threshold, and the approximate posterior were exact then the distribution of FPs is binomial with rate $1/N$. The expected number of false positives in each PPM is therefore $N \times 1/N = 1$. The variance is $N \times 1/N \times (1 - 1/N)$ which is approximately 1. We would therefore expect 0, 1 or 2 false positives per PPM.

Of course, the above result only holds if the approximate posterior is equal to the true posterior. But given that all of our computational effort is aimed at this goal it would not be surprising if the above analysis were indicative of actual FP rates. This issue will be investigated using null fMRI data in the next section.

Results

Null data

This section describes the analysis of a Null data set to find out how many false positives are obtained using PPMs with default thresholds.

Images were acquired from a 1.5T Sonata (Siemens, Erlangen Germany) which produced T2*-weighted transverse Echo-Planar Images (EPIs) with BOLD contrast, whilst a subject was lying in the scanner, asked to rest and was not provided with any experimental stimulus. These data are thus collected under the null hypothesis, H_0, that experimental effects are zero. This should hold whatever the design matrix and contrast we conceive. Any observed activations will be false positives.

Whole brain EPIs consisting of 48 transverse slices were acquired every TR=4.32s resulting in a total of $T = 98$ scans. The voxel size is $3 \times 3 \times 3$mm. All images were realigned to the first image using a six-parameter rigid-body transformation to account for subject movement. These data were not spatially smoothed. Whilst spatial smoothing is necessary for the standard application of classical inference (see eg. chapter 2), it is not necessary for the spatio-temporal models described in this chapter. Indeed, the whole point of SRGLM is that the optimal smoothness can be inferred from the data.

We then implemented a standard whole volume analysis on images comprising $N = 59,945$ voxels. We used the design matrix shown in the left panel of Figure 2. Use of the default thresholds resulted in no spurious activations in the PPM.

We then repeated the above analysis but with a number of different design matrices. Firstly, we created a number of epoch designs. These were based on the design in Figure 2 but epoch onsets were jittered by a number between plus or minus 9 scans. This number was drawn from a uniform distribution, and the epoch durations were drawn from a uniform distribution between 4 and 10 scans. Five such design matrices were created and VB models fitted with each to the null data. For every analysis, the number of false positives was 0.

Secondly, we created a number of event-related designs by sampling event onsets with inter-stimulus intervals drawn from a poisson distribution with rate 5 scans. These event streams were then convolved with a canonical HRF [Henson 2003]. Again, five such design matrices were created and VB models fitted with each to the null data. Over the 5 analyses, the average number of false positives was 9.4. The higher false positive rate for event-related designs is thought to occur because event-related regressors are more similar than epoch regressors to fMRI noise.

Synthetic data

We then added three synthetic activations to a slice of null data ($z = -13$mm). These were created using the design matrix and regression coefficient image shown in Figure 2 (the two regression coefficient images, ie. for the activation and the mean, were identical). These images were formed by placing delta functions at three locations and then smoothing with Gaussian kernels having FWHMs of 2, 3 and 4 pixels (going clockwise from the top-left blob). Images were then rescaled to make the peaks unity.

In principle, smoothing with a Gaussian kernel renders the true effect size greater than zero everywhere because a Gaussian has infinite spatial support. In practice, however, when implemented on a digital computer with finite numerical precision most voxels will be numerically zero. Indeed, our simulated data contained 299 'activated' voxels ie. voxels with effect sizes numerically greater than zero.

This slice of data was then analysed using VB. The contrast $C^T = [1, 0]$ was then used to look at the estimated activation effect which is shown in the left panel of Figure 3. For comparison, we also show the effect as estimated using OLS. Clearly, OLS estimates are much noisier than VB estimates.

Figure 4 shows plots of the exceedance probabilities for two different effect-size thresholds, $\gamma = 0$ and $\gamma = 0.3$. Figure 5 shows thresholded versions of these images. These are PPMs. Neither of these PPMs contain any false positives. That is, the true effect size is greater than zero wherever a white voxel occurs. This shows, informally, that use of the default thresholds provides good specificity whilst retaining reasonable sensitivity. Also, a combination of non-zero effect-size thresholds and more liberal probability thresholds can do the same.

Auditory data

This section describes the use of multivariate contrasts for an auditory fMRI data set comprising whole brain BOLD/EPI images acquired on a modified 2T Siemens Vision system. Each acquisition consisted of 64 contiguous slices ($64 \times 64 \times 64$, 3mm $\times$ 3mm $\times$ 3mm voxels) and a time series of 84 images was acquired with TR=7s from a single subject.

This was an epoch fMRI experiment in which the condition for successive epochs alternated between rest and auditory stimulation, starting with rest. Auditory stimulation was bi-syllabic words presented binaurally at a rate of 60 per minute.

These data were analysed using VB with the design matrix shown in Figure 6. To look for voxels that increase activity in response to auditory stimulation we

used the univariate contrast $C^T = [1, 0]$. Figure 7 shows a PPM that maps effect-sizes of above threshold voxels.

To look for either increases or decreases in activity we use the multivariate contrast $C^T = [1, 0]$. This inference uses the χ^2 approach described earlier. Figures 8 shows the PPM obtained using default thresholds.

Face data

This is an event-related fMRI data set acquired by Henson et al. [Henson et al. 2002] during an experiment concerned with the processing of faces. Greyscale images of faces were presented for 500ms, replacing a baseline of an oval chequerboard which was present throughout the interstimulus interval. Some faces were of famous people, and were therefore familiar to the subject, and others were not. Each face in the database was presented twice. This paradigm is a two-by-two factorial design where the factors are familiarity and repetition. The four experimental conditions are 'U1', 'U2', 'F1' and 'F2' which are the first or second (1/2) presentations of images of familiar 'F' or unfamiliar 'U' faces.

Images were acquired from a 2T VISION system (Siemens, Erlangen, Germany) which produced T2*-weighted transverse Echo-Planar Images (EPIs) with BOLD contrast. Whole brain EPIs consisting of 24 transverse slices were acquired every two seconds resulting in a total of T=351 scans. All functional images were realigned to the first functional image using a six-parameter rigid-body transformation. To correct for the fact that different slices were acquired at different times, time series were interpolated to the acquisition time of a reference slice. Images were then spatially normalized to a standard EPI template using a nonlinear warping method [Ashburner and Friston 2003]. Each time series was then high-pass filtered using a set of discrete cosine basis functions with a filter cut-off of 128 seconds.

The data were then analysed using the design matrix shown in Figure 9. The first 8 columns contain stimulus related regressors. These correspond to the four experimental conditions, where each stimulus train has been convolved with two different hemodynamic bases (i) the canonical Hemodynamic Response Function (HRF) and (ii) the time derivative of the canonical [Henson 2003]. The next 6 regressors in the design matrix describe movement of the subject in the scanner and the final column models the mean response.

Figure 10 plots a map of the first autoregressive component as estimated using VB. This shows a good deal of heterogeneity and justifies our assumption that that AR coefficients are spatially varying. The estimated spatial variation is smooth, however, due to the spatial prior. Figure 11 shows a PPM for 'Any effect of faces' which was obtained using the multivariate contrast matrix shown in Figure 9.

Discussion

We have reviewed a framework for the analysis of fMRI data based on spatially regularised GLMs. This model embodies prior knowledge that evoked responses are spatially homogeneous and locally contiguous.

As compared to standard approaches based on spatially smoothing the imaging data itself, the spatial regularisation procedure has been shown to result in

inferences with higher sensitivity [Penny et al. 2005c]. The approach may be viewed as an automatically regularized spatio-temporal deconvolution scheme.

Use of PPMs with default thresholds resulted in low false positive rates for null fMRI data, and physiologically plausible activations for auditory and face fMRI data sets. We have recently developed a similar approach for source localisation of EEG/MEG, which is described in the following chapter.

References

[SPM Software 2006] SPM Software, Wellcome Department of Imaging Neuroscience. Available from http://www.fil.ion.ucl.ac.uk/spm/software/, 2006.

[Ashburner and Friston 2003] J. Ashburner and K.J. Friston. Spatial normalization using basis functions. In R.S.J. Frackowiak, K.J. Friston, C. Frith, R. Dolan, K.J. Friston, C.J. Price, S. Zeki, J. Ashburner, and W.D. Penny, editors, *Human Brain Function*. Academic Press, 2nd edition, 2003.

[Auranen 2005] T. Auranen, A. Nummenmaa, M. Hammalainen, I. Jaaskelainen, J. Lampinen, A. Vehtari, and M. Sams. Bayesian analysis of the neuromagnetic inverse problem with l^p norm priors. *Neuroimage*, 26(3):870–884, 2005.

[Box and Tiao 1992] G.E.P Box and G.C. Tiao. *Bayesian Inference in Statistical Analysis*. John Wiley, 1992.

[Frackowiak et al. 2003] R.S.J. Frackowiak, K.J. Friston, C. Frith, R. Dolan, C.J. Price, S. Zeki, J. Ashburner, and W.D. Penny. *Human Brain Function*. Academic Press, 2nd edition, 2003.

[Friston et al. 1995] K.J. Friston, A.P. Holmes, K.J. Worsley, J.B. Poline, C. Frith, and R.S.J. Frackowiak. Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2:189–210, 1995.

[Friston and Penny 2003] K.J. Friston and W.D. Penny. Posterior probability maps and SPMs. *NeuroImage*, 19(3):1240–1249, 2003.

[Gautama and Van Hulle 2004] T. Gautama and M.M. Van Hulle. Optimal spatial regularisation of autocorrelation estimates in fMRI analysis. *NeuroImage*, (23):1203–1216, 2004.

[Harrison et al. 2003] L. Harrison, W.D. Penny, and K.J. Friston. Multivariate autoregressive modelling of fMRI time series. *NeuroImage*, 19(4):1477–1491, 2003.

[Henson 2003] R.N.A. Henson. Analysis of fMRI time series. In R.S.J. Frackowiak, K.J. Friston, C. Frith, R. Dolan, K.J. Friston, C.J. Price, S. Zeki, J. Ashburner, and W.D. Penny, editors, *Human Brain Function*. Academic Press, 2nd edition, 2003.

[Henson et al. 2002] R.N.A. Henson, T. Shallice, M.L. Gorno-Tempini, and R.J. Dolan. Face repetition effects in implicit and explicit memory tests as measured by fMRI. *Cerebral Cortex*, 12:178–186, 2002.

[Pascal Marqui et al. 1994] R. Pascual Marqui, C. Michel, and D. Lehman. Low resolution electromagnetic tomography: a new method for localizing electrical activity of the brain. *International Journal of Psychophysiology*, pages 49–65, 1994.

[Penny et al. 2003] W.D. Penny, S.J. Kiebel, and K.J. Friston. Variational Bayesian Inference for fMRI time series. *NeuroImage*, 19(3):727–741, 2003.

[Penny and Flandin 2005] W.D. Penny and G. Flandin. Bayesian analysis of fMRI data with spatial priors. In *Proceedings of the Joint Statistical Meeting (JSM), Section on Statistical Graphics [CDROM]*, 2005. Alexandria, VA: American Statistical Association.

[Penny and Flandin 2005b] W.D. Penny and G. Flandin. Bayesian analysis of single-subject fMRI: SPM implementation. Technical report, Wellcome Department of Imaging Neuroscience, 2005.

[Penny et al. 2005c] W.D. Penny, N. Trujillo-Barreto, and K.J. Friston. Bayesian fMRI time series analysis with spatial priors. *NeuroImage*, 24(2):350–362, 2005.

[Penny et al. 2006] W.D. Penny, G. Flandin, and N. Trujillo-Barreto. Bayesian Comparison of Spatially Regularised General Linear Models. *Human Brain Mapping*, 2006. Accepted for publication.

[Woolrich et al. 2001] M.W. Woolrich, B.D. Ripley, M. Brady, and S.M. Smith. Temporal autocorrelation in univariate linear modelling of fMRI data. *NeuroImage*, 14(6):1370–1386, December 2001.

[Woolrich et al. 2004] M.W. Woolrich, T.E. Behrens, and S.M. Smith. Constrained linear basis sets for HRF modelling using Variational Bayes. *NeuroImage*, 21:1748–1761, 2004.

Appendix

This appendix provides a number of formulae required for updating the approximate posteriors. These have been derived in [Penny et al. 2003]. Firstly, auxiliary quantities for updating $q(w_n)$ are

$$\tilde{A}_n = \sum_t x_t^T x_t + X_t^T (m_n^T m_n + V_n) X_t - x_t^T m_n X_t - X_t^T m_n^T x_t \tag{24}$$

$$\tilde{b}_n = \sum_t y_{tn} x_t - m_n d_{tn} x_t - y_{tn} m_n X_t + d_{tn}^T (m_n^T m_n + V_n) X_t \tag{25}$$

For the special case in which the errors are uncorrelated, ie. $P = 0$, we have $\tilde{A}_n = X^T X$ and $\tilde{b}_n = X^T y_n$. If we also have no spatial prior on the regression coefficients, ie. $\alpha = 0$, we then recover the least squares update

$$\hat{w}_n \quad = \quad (X^T X)^{-1} X^T y_n \tag{26}$$

Secondly, auxiliary quantities for updating $q(a_n)$ are

$$\begin{aligned}\tilde{C}_n &= \sum_t d_{tn} d_{tn}^T + X_t(\hat{w}_n \hat{w}_n^T + \hat{\Sigma}_n) X_t^T \\ &- d_{tn} \hat{w}_n^T X_t^T - X_t \hat{w}_n d_{tn}^T \\ \tilde{d}_n &= \sum_t y_{tn} d_{tn}^T - x_t \hat{w}_n d_{tn}^T - y_{tn} \hat{w}_n^T \tilde{X}^T + x_t(\hat{w}_n \hat{w}_n^T + \hat{\Sigma}_n) X_t^T\end{aligned} \tag{27}$$

Thirdly, the auxiliary quantity for updating $q(\lambda_n)$ is

$$\tilde{G}_n = \tilde{G}_{n1} + \tilde{G}_{n2} + \tilde{G}_{n3} \tag{28}$$

where

$$\begin{aligned}\tilde{G}_{n1} &= \sum_t y_{tn}^2 + d_{tn}^T (m_n^T m_n + V_n) d_{tn} - 2 y_{tn} d_{tn}^T m_n \\ \tilde{G}_{n2} &= \sum_t x_t(\hat{w}_n \hat{w}_n^T + \hat{\Sigma}) x_t^T + \mathsf{Tr}(X_t^T (m_n^T m_n + V_n) X_t \hat{\Sigma}_n) \\ &+ \hat{w}_n^T X_t^T (m_n^T m_n + V_n) X_t \hat{w}_n - 2 x_t (\hat{w}_n \hat{w}_n^T + \hat{\Sigma}_n) X_t m_n^T \\ \tilde{G}_{n3} &= \sum_t -2 y_{tn} x_t \hat{w}_n + 2 m_n d_{tn} x_t \hat{w}_n \\ &+ 2 y_{tn} m_n X_t \hat{w}_n - 2 d_{tn}^T (m_n^T m_n + V_n) X_t \hat{w}_n\end{aligned} \tag{29}$$

Many terms in the above equations do not depend on model parameters and so can be pre-computed for efficient implementation. See [Penny and Flandin 2005b] for more details.

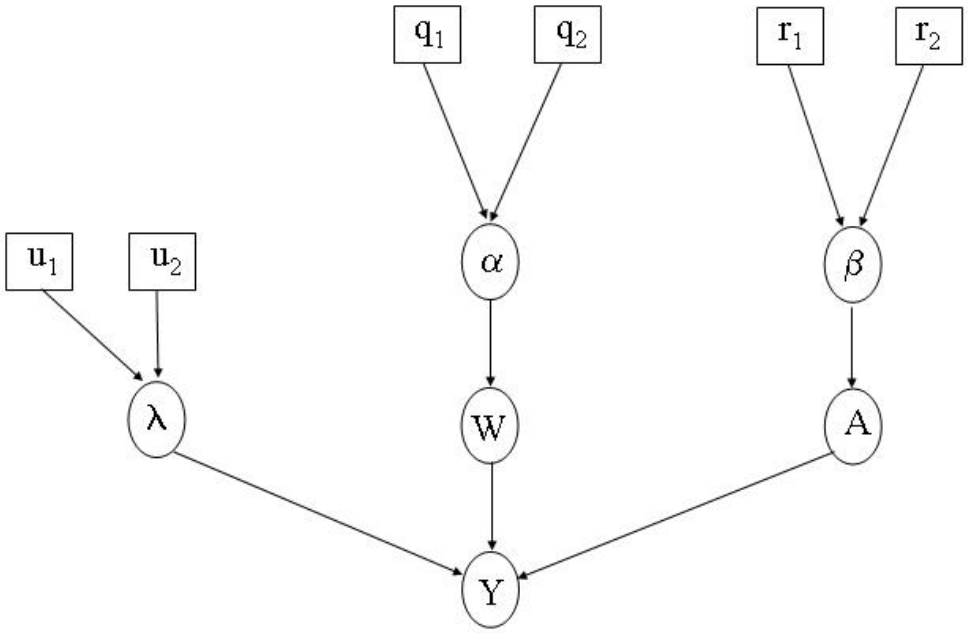

Figure 1: *The figure shows the probabilistic dependencies underlying the SRGLM generative model for fMRI data. The quantities in squares are constants and those in circles are random variables. The spatial regularisation coefficients α constrain the regression coefficients W. The parameters λ and A define the autoregressive error processes which contribute to the measurements. The spatial regularisation coefficients β constrain the AR coefficients A.*

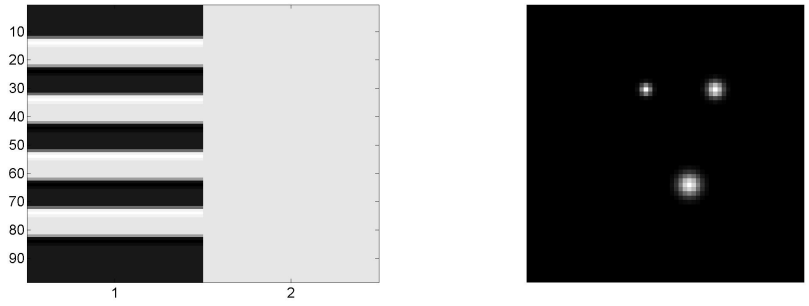

Figure 2: *Left: Design matrix for null fMRI data. The first column models a boxcar activation and the second column models the mean. There are* $n = 1..98$ *rows corresponding to the 98 scans. Right: Image of regression coefficients corresponding to a synthetic activation. This image is added to the null data. In this image, and others that follow, black is 0 and white is 1.*

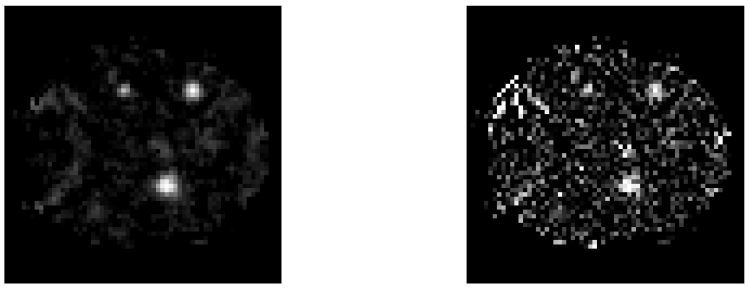

Figure 3: *Left: Estimated effect using VB (the true effect is shown in the right plot in Figure 2). Right: Estimated effect using OLS.*

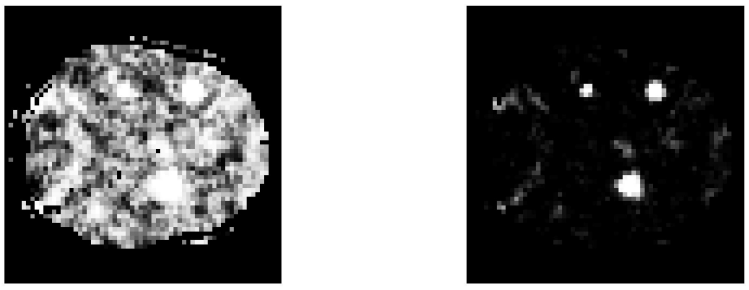

Figure 4: *Plots of exceedance probabilities for two* γ *thresholds. Left: A plot of* $p(c_n > 0)$*. Right: A plot of* $p(c_n > 0.3)$*.*

Figure 5: *PPMs for two thresholds. Left: The default thresholds ($\gamma = 0$, $p_T = 1 - 1/N$) Right: The thresholds $\gamma = 0.3$, $p_T = 0.95$.*

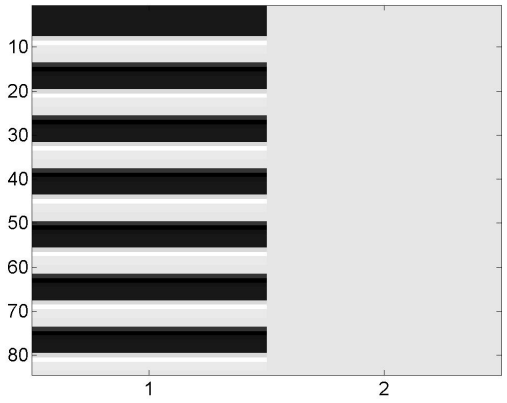

Figure 6: *Design matrix for analysis of the auditory data. The first column models epochs of auditory stimulation and the second models the mean response.*

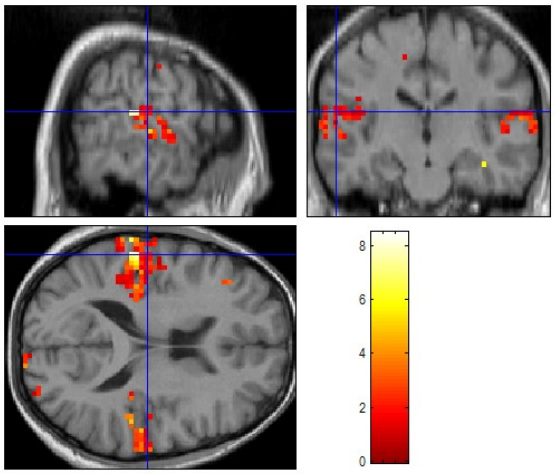

Figure 7: *PPM for positive auditory activation. Overlay of effect-size, in units of percentage of global mean, on subjects MRI for above threshold voxels. The default thresholds were used, that is, we plot* c_n *for voxels which satisfy* $p(c_n > 0) > 1 - 1/N$.

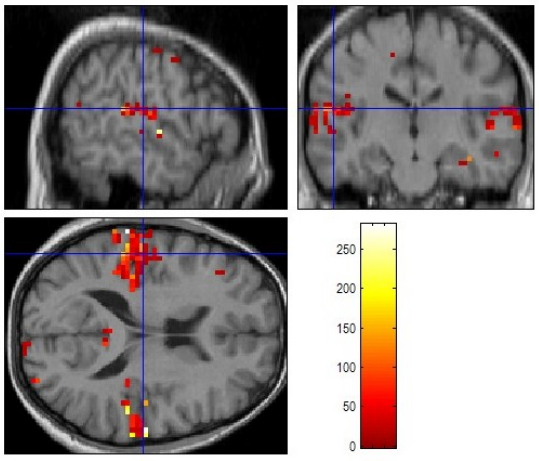

Figure 8: *PPM for positive or negative auditory activation. Overlay of* χ^2 *statistic on subjects MRI for above threshold voxels. The default thresholds were used, that is, we plot* χ^2_n *for voxels which satisfy* $p(c_n > 0) > 1 - 1/N$.

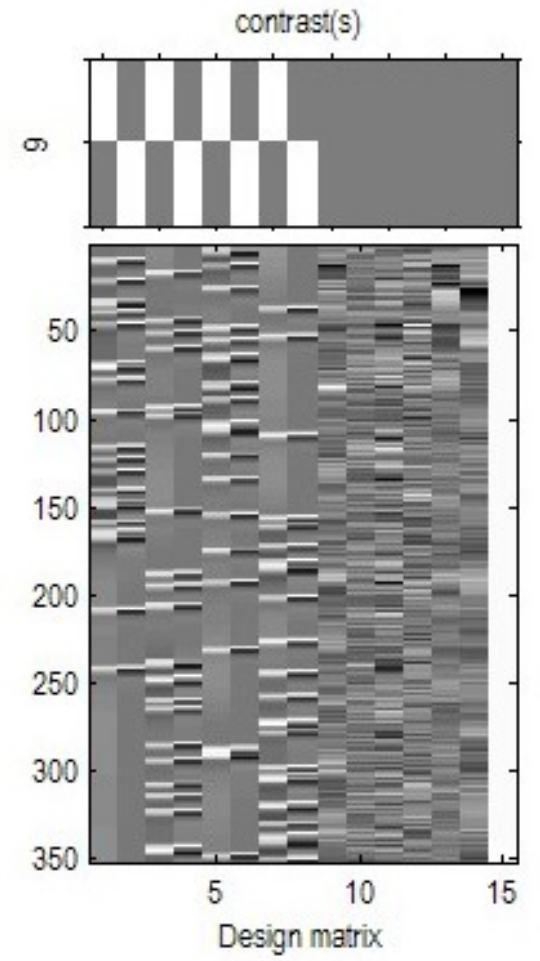

Figure 9: *Lower part: Design matrix for analysis of face data, Upper part: Multivariate contrast used to test for any effect of faces.*

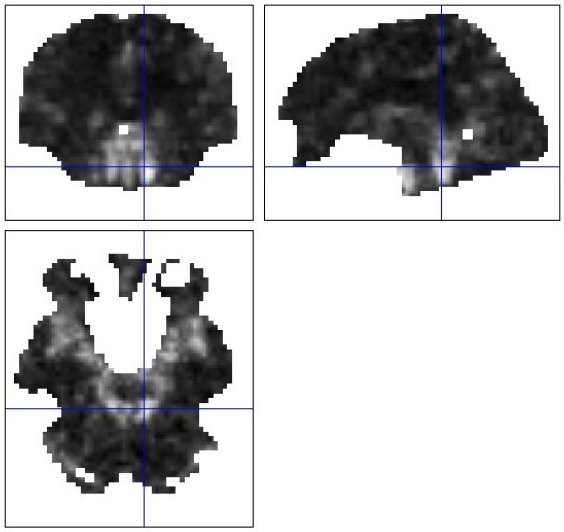

Figure 10: *Image of the first autoregressive coefficient estimated from the face fMRI data (in all, there were $P = 3$ AR coefficients per voxel). Black denotes 0 and white 1. Large values can be seen around the circle of Willis and middle cerebral artery. This makes sense as cardiac-induced pulsatile motion is likely to be stronger in these regions.*

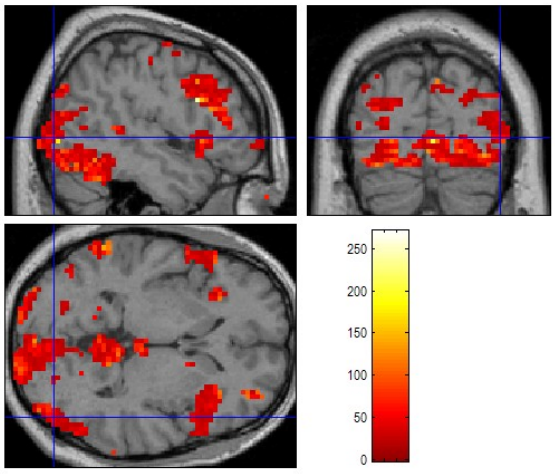

Figure 11: *PPM showing above threshold χ^2 statistics for any effect of faces*