Characterizing Dynamic Brain Responses with fMRI: A Multivariate Approach

K. J. Friston, C. D. Frith, R. S. J. Frackowiak, and R. Turner

The Wellcome Department of Cognitive Neurology, Queen Square, WC1N 3BG United Kingdom; and the MRC Cyclotron Unit, Hammersmith Hospital W12 OHS United Kingdom

Received April 17, 1995

In this paper we present a multivariate analysis of evoked hemodynamic responses and their spatiotemporal dynamics as measured with fast fMRI. This analysis uses standard multivariate statistics (MANCOVA) and the general linear model to make inferences about effects of interest and canonical variates analysis (CVA) to describe the important features of these effects. We have used these techniques to characterize the form of hemodynamic transients that are evoked during a cognitive or sensorimotor task. In particular we do not assume that the neural or hemodynamic response reaches some "steady state" but acknowledge that these physiological changes could show profound task-dependent adaptation and time-dependent changes during the task. To address this issue we have modeled hemodynamic responses using appropriate temporal basis functions and estimated their exact form within the general linear model using MANCOVA. We do not propose that this analysis is a particularly powerful way to make inferences about functional specialization (or more generally functional anatomy) because it only provides statistical inferences about the distributed (whole brain) responses evoked by different conditions. However, its application to characterizing the temporal aspects of evoked hemodynamic responses reveals some compelling and somewhat unexpected perspectives on transient but stereotyped responses to changes in cognitive or sensorimotor processing. The most remarkable observation is that these responses can be biphasic and show profound differences in their form depending on the extant task or condition. Furthermore these differences can be seen in the absence of changes in mean signal.

INTRODUCTION

Fast fMRI techniques allow us to look at evoked physiological responses, in the brain, on a second by second basis. However, the technique has been used primarily to assess hemodynamic indices while assuming that the physiological concomitants of cognitive or sensorimotor processing are in some "steady state." This begs the question: "Why acquire data at such a fast rate?" In this short paper, and a companion paper (Friston et al., 1995a), we present compelling evidence that suggests that it is potentially important to look at brain changes on a second by second basis. Our main result is that the shape and duration of hemodynamic transients, which ensue in the seconds following a change in task, can vary with the task, and furthermore, the form of these transients can discriminate between tasks even if the average signal is the same for all tasks. The evidence for this conclusion is based on a multivariate analysis applied to a fMRI time-series obtained during a motor sequencing activation study reported in this paper. The results were then used to inform and extend our current univariate approach to increase its sensitivity by modeling hemodynamic transients in an empirically directed and more appropriate way. This extension is described in a companion paper.

The problem addressed by this paper is how to assess and characterize evoked spatiotemporal hemodynamics changes using fast fMRI. Current approaches to this problem overlook the fact that, following the onset of an "activation," the brain responds with a rich and protracted hemodynamic transient, with many variations over both space and time. Even approaches that explicitly refer to a nonlinear hemodynamic response function (e.g., Friston et al., 1994, 1995b) assume that the underlying neural dynamics are unchanging during an activation. This assumption is implicit in the square-wave or box-car reference vectors used in the analyses. In this paper we adopt a different perspective, borrowed from the field of evoked potentials using EEG and MEG. This view considers that the evoked response is a stereotyped transient of unknown form that can change from brain region to brain region and from task to task. By using appropriate temporal basis functions it is possible to estimate the form of these
transients and make some statistical inferences about them. Specifically, we introduce multivariate analysis of variance (MANCOVA) and canonical variates analysis (CVA) to characterize evoked hemodynamic responses in space and time. The importance of this analysis is fourfold: (i) Unlike existing approaches it provides for statistical inferences about the significance of the spatiotemporal response over the entire volume. (ii) The approach implicitly takes account of spatial correlations in the data without making any assumptions. (iii) The canonical variates analysis produces de-noised eigenimages that characterize the activation effects in terms of a series of spatial modes (canonical images) and (canonical) transients. (iv) The theoretical basis is well established and can be found in any introductory text on multivariate analysis.

Unlike statistical parametric mapping with univariate tests (e.g., t maps and Z maps), MANCOVA is explicitly multivariate. It considers one observation as comprising all voxels in a scan. In other words, one response variable corresponds to all its components (voxel values) in the volume. The importance of this multivariate approach is that the effects due to activations, confounding effects, and error effects are assessed in terms of both the effects at each voxel and interactions among voxels. This means that one does not have to model spatial correlations (e.g., with Gaussian Fields) when assessing the significance of an activation effect. These correlations are explicitly included in the analysis. The MANCOVA uses the general linear model to model effects of interest and provides a single P value reflecting the significance of these effects. In this application the effects of interest are some systematic response that endures during each condition, but whose form is unknown. These transients are modeled using a small set of Fourier basis functions such that the estimated transients can have any (smooth) shape.

Having established significance using MANCOVA, the nature of the spatiotemporal response remains to be characterized. We propose that Canonical Variates Analysis (CVA) is an appropriate characterization of distributed activation effects. Canonical images are similar to eigenimages but derive from a statistical model with error terms (the general linear model used by the MANCOVA). CVA is closely related to denoising techniques in EEG and MEG time series analysis that use a generalized eigenvalue solution. Intuitively these approaches can be understood as finding the eigenimages that “point toward the activation effects and away from the noise” (Anders Dale, personal communication). Another way of looking at canonical images, obtained with CVA, is to think of them as statistically informed eigenimages that discount interactions due to error.

The paper is divided into two sections. The first section deals with the theory behind MANCOVA and CVA. It presents the operational equations behind the multivariate general linear model, statistical inference about activation effects based on Wilk’s Lambda and the canonical values, and characterizing the nature of these effects using CVA. The second section describes an application to a fMRI activation study of attention and motor sequencing. The primary aim of this analysis was to characterize the nature and form of transient hemodynamic responses to changing task conditions as a prelude to improving existing univariate techniques. We reiterate that the procedures described in this paper can be found in any standard introductory test on multivariate statistics. We have used Chatfield and Collins (1980).

**THEORY**

The terminology adopted in this paper divides a fMRI time-series of (I) scans into meaningful blocks called “epochs” that correspond to the presence of a particular task or condition. A single multivariate response variable contains all the voxel values in one scan. The first step in multivariate analysis is to ensure that the dimensionality of the data is smaller than the number of observations (I). Clearly, this is not the case because there are more voxels than scans; therefore, the data have to be transformed. The dimension reduction proposed here is straightforward and uses the eigenvector solution of the sums of squares and products over time to give a reduced set of components for each multivariate observation. These eigenimages and their expression in time (eigenvectors) can be calculated in a number of ways. We use the standard eigenvalue solution in this paper:

\[
[\epsilon \lambda] = \text{eig}(XX^T),
\]

where

\[
(XX^T) \cdot \epsilon = \epsilon \cdot \lambda
\]

\[
U = XX^T \cdot \epsilon \cdot \lambda^{-\frac{1}{2}}
\]

and

\[
X = \epsilon \cdot \lambda^2.
\]  

(1)

Here \(X^T\) is a large matrix of corrected voxel values with one column for each voxel and one row for each scan. By “corrected” we mean that obvious confounds such as low frequency artifacts and global effects have been removed (using linear regression) and that the data have been mean corrected. \(\lambda\) is a diagonal matrix of
eigenvalues and $\varepsilon$ is a matrix of eigenvectors over time. The spatial modes or eigenimages correspond to the columns of $\mathbf{U}$ and their expression over time to the columns of $\mathbf{X}$. $\mathbf{X}$ has one column for (at most) every eigenimage with a nonzero eigenvalue and one row for each scan. In our work we use only the columns of $\mathbf{X}$ that have an associated eigenvalue greater than unity. We present the derivation of the eigenimages in this rather clumsy way because computationally it is much easier to compute $\text{eig} (\mathbf{X}^* \mathbf{X}^*)$ than $\text{eig} (\mathbf{X}^* \mathbf{X}^*)$. Intuitively $\mathbf{X}$ can be thought of as the original data $\mathbf{X}^*$ "looked at" from a different direction. The element of $\mathbf{X}$ are $x_{ij}$ the activity of the $j$th eigenimage in scan $i$.

In matrix notation the general linear model is

$$\mathbf{X} = \mathbf{G}\beta + \varepsilon$$

(2)

The general linear model assumes the errors $\varepsilon$ are independent and identically distributed with the normal distribution $[\mathcal{N}(0, \sigma^2_{\varepsilon})]$. If the data have been smoothed in time then a correction to the degrees of freedom associated with the error terms is required (see Friston et al., 1995b, for more details). The matrix $\mathbf{G}$ is called the design matrix. The design matrix has one column for every effect (factor or covariate) in the model. $\beta$ is the parameter matrix with one column vector $\beta_j$ of parameters for each eigenimage. The elements of $\mathbf{G}$ are explanatory variables relating to the conditions under which the observation (e.g., scan) was made. In this application the design matrix contains a small set of Fourier basis functions for each epoch. This set is shown in Fig. 1. The design matrix is constructed such that these four basis functions are repeated for all instances of a particular condition. For example, if we had three conditions the design matrix would have 12 columns. Each set of four columns would contain the four basis functions whenever that condition occurred (see below). Least squares estimates of the contribution of these basis functions to each condition are given by the estimates of $\beta$ say $\mathbf{b}$:

$$\mathbf{b} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{X}$$

(3)

and, if the error terms are independent

$$\text{Var} [\mathbf{b}_j] = \sigma^2 (\mathbf{G}^T \mathbf{G})^{-1}.$$  

(4)

### Statistical Inference

In this section we address statistical inference about the effects of interest (condition or covariates of interest). Significance is assessed by testing the null hypothesis that the effects of interest do not significantly reduce the error variance (or alternatively the null hypothesis that $\beta$ is zero). The null hypothesis can be tested in the following way. The sum of squares and products due to error $\mathbf{R}(\Omega)$ is obtained from the difference between the actual and estimated values of $\mathbf{X}$:

$$\mathbf{R} = \mathbf{R}(\Omega) = (\mathbf{X} - \mathbf{G} \mathbf{b})^T (\mathbf{X} - \mathbf{G} \mathbf{b}),$$

(5)

where the sums of squares and products due to effects of interest are given by:

$$\mathbf{T} = (\mathbf{G} \mathbf{b})^T (\mathbf{G} \mathbf{b}).$$

(6)

The error sum of squares and products under the null hypothesis $\mathbf{R}(\Omega_0)$, i.e., after discounting the effects of interest ($\mathbf{G}$) is simply

$$\mathbf{R}(\Omega_0) = \mathbf{X}^T \mathbf{X} = \lambda.$$

(7)

The significance can now be tested with

$$\Lambda = |\mathbf{R}(\Omega)| / |\mathbf{R}(\Omega_0)|,$$

(8)

where $\Lambda$ is Wilk's statistic (known as Wilk's Lambda). Under the null hypothesis, and after transformation $\Lambda$ has a $\chi^2$ distribution:

$$-(r - ((J - h + 1)/2)) \log(\Lambda) \sim \chi^2(J, h),$$

(9)

where $r$ is the degrees of freedom associated with the error terms. We use the effective degrees of freedom.
computed as described in Friston et al. (1995b) \( r = \frac{\text{rank}(X^* X^{\top}) - \text{rank}(G)}{\sqrt{2\pi s^2}} \) and \( s \) is the standard deviation of any Gaussian kernel used to smooth the data in time. \( J \) is the number of eigenimages (e.g., with eigenvalues greater than 1) used in the \( J \)-ivariate response variable \( X \) and \( h \) is the degrees of freedom associated with the effects of interest = rank\( (G) \).

**Characterizing the Effect**

Having established that the effects of interest are significant (e.g., differences among the task-related epochs) the final step is to characterize these effects in terms of their spatiotemporal dynamics. This characterization uses CVA. The objective is to find the linear combination (compound or contrast) of the components of \( X \), in this case the eigenimages, that best express the activation effects when compared to error effects. More exactly we want to find \( c_2 \) such that the variance ratio

\[
(c_1^T \cdot T \cdot c_1) / (c_1^T \cdot R \cdot c_1)
\]

is maximized. Let \( Z_1 = X \cdot c_1 \), where \( Z_1 \) is the first canonical variate and \( c_1 \) is a canonical vector that maximizes this ratio. \( c_2 \) is the second canonical vector that maximizes the ratio subject to the constraints the Cov\( (c_1, c_2) = 0 \) (and so on). The matrix of canonical images \( C = [c_1, c_2, \ldots, c_J] \) is given by solution of the generalized eigenvalue problem

\[
T \cdot c = R \cdot c \cdot \theta,
\]

where \( \theta \) is a diagonal matrix of scaled canonical values. Canonical images \( C \) are obtained by rotating the canonical vector in the columns of \( C \) back into “voxel” space with the original eigenimages \( U \):

\[
C = U \cdot C.
\]

The columns of \( C \) now contain the voxel values of the canonical images. The \( j \)th column of \( C \) (the \( j \)th canonical image) has an associated canonical value equal to \( j \)th leading diagonal element of \( \theta \). Note that the “activation” effect is a multivariate one, with \( J \) components. The final step involves determining the number of canonical images that can be considered “significant.” This is effected by testing for the dimensionality of the response using the canonical values \( \theta \), where under the null hypothesis the probability that the dimensionality of the response is greater than \( S \) is tested with

\[
(r - ((J - h + 1)/2)) \cdot \log \left( \prod_{j=S+1}^{J} (1 + \theta_j) \right) - \chi^2((J - S) \cdot (h - S))
\]

**AN EMPIRICAL APPLICATION**

In these sections we present an example of the analysis. The example chosen is typical of more sophisticated fMRI activation studies that use more than two conditions. There were three conditions: a rest condition, a motor sequencing condition where the subject moved his or her right and left hand in a fixed alternating order (in response to a visual cue), and a motor sequencing condition in which the subject moved either the right or the left hand as instructed visually in a random sequence. The only difference between the “fixed” and “random” conditions was that the subject had to attend to the instruction that specified the movement. Clearly in the random condition the subject could not anticipate or prepare the exact movement before seeing the cue. The movements involved raising the forefinger and the cues were presented at pseudo-random intervals of 2.3 or 4. s. The details of this paradigm and a full discussion of the results will be presented elsewhere.

**Data Acquisition**

One hundred and twenty \( T_2 \) weighted volume images \((128 \times 64 \times 10 \text{ voxels})\) were obtained from a single male subject using a GE/ANMR 1.5T system with EPI capabilities. The volumes consisted of 10 sequential transverse sections and were acquired every 3 s. Voxel size was \( 3 \times 3 \times 7 \text{ mm} \) voxels with 0.5-mm slice separation. The three conditions were presented in blocks of 10 scans, in pseudorandom order. Each condition was therefore repeated four times, each time constituting a 30-s epoch for that condition.

**Data Prepossessing**

The 120 volume images were realigned to the first as described elsewhere (Friston et al., 1995c) and resampled to \( 3 \times 3 \times 6 \text{ mm} \) voxels. The data were then smoothed with an isotropic Gaussian kernel with FWHM of 8 mm in space (in this analysis we were more interested in spatially distributed temporal dynamics and therefore used a larger spatial filter than is typical in fMRI) and \( \sqrt{5} \text{ s} \) in time. Voxels that had values greater than 0.8 of the volume mean in all the images were selected to restrict the analysis to intracranial regions. The confounding effects of global (whole volume) activity and time and were removed using linear regression with global activity and sine/cosine functions.
as confounds (up to a maximum of 2.5 cycles per 120 scans). Removing the latter confounds corresponds to high-pass filtering the time-series to remove low frequency artifacts due to aliased cardiorespiratory and other cyclical components. The data were mean corrected to form the matrix \( X^* \) above.

**MANCOVA**

The corrected data were reduced to 35 eigenvectors using Eq. (1) (i.e., there were 35 eigenvalues greater than unity) and subject to MANCOVA as described in the theory section. The design matrix \( G \) is shown, in image format, in the upper panel of Fig. 2. The first four columns correspond to the rest conditions, the second four to the fixed conditions, and the last four to the random conditions. The significance of the condition dependent effects was assessed with Wilk's Lambda which was 620.16. Using Eq. (8) the corresponding \( P \) value for the activation effects was \( P < 10^{-8} \). The degrees of freedom due to error were 54 and the degrees of freedom for Wilk's Lambda were \( 12 \times 35 = 420 \).

**CVA**

The spectrum of canonical values \( \theta_i \) is shown in the lower left panel of Fig. 2 and the corresponding \( P \) values are seen on the lower right. It can be seen that the first three \( P \) values are less than 0.05 and therefore the dimensionality of the response is threefold. In other words the spatiotemporal dynamical response to these changing conditions requires at least three spatial modes to describe it. In this paper we focus on the first or most important canonical image. A maximum intensity projection of the first canonical image (positive components) and its time-dependent expression or canonical variate \( Z_1 = Xc_1 = X^*c_1 \) are shown in Fig. 3. The upper panel demonstrates that this system is distributed with a rather diffuse spatial topography. The spatial characteristics of this mode should not be over-

**FIG. 2.** (Top left) Design matrix: This is an image representation of the design matrix \( G \). Because elements of this matrix can take negative values the gray scale is arbitrary and has been scaled to the minimum and maximum. (Bottom left) Spectrum of canonical values \( \theta \) following a canonical variates analysis (Bottom right) Spectrum of \( P \) values based on \( \theta \) the roots of \( R^{-1}T \). These \( P \) values test the hypothesis that the dimensionality of the result is greater than a specified number \( (S) \) and are based on the distributional approximation given in Eq. (12).

**FIG. 3.** (Top left) First canonical image: This is a maximum intensity projection of positive parts of the first canonical image. The display format is standard and provides three views of the brain from the front, the bottom, and the left-hand side. The gray scale is arbitrary and the space conforms to that described in the atlas of Talairach and Tournoux (1988). (Bottom panel) Canonical variates expressed as a function of time (in scans). The dotted line in the left panel corresponds to the first canonical variate \( Z_1 = Xc_1 \) and the solid line to \( Gc_1 \). This representation of the canonical variate is shown again on the right for a single epoch of each condition.
interpreted (this canonical image simply reflects the amount of variance that each region contributes to the first canonical variate. This amount may be small or large relative to other components). However, it is worth noting that some regions show greater loadings than others, in particular, the anterior prefrontal regions, the SMA, and parietal regions.

Of greater interest is the canonical variate shown in the lower left panel of Fig. 3. The dotted line corresponds to $Z_1$ and the solid line to the canonical variate expressed in terms of the modeled effects (i.e., $G \cdot b \cdot c_1$). The solid line is therefore a succession of canonical transients expressed in response to each of the three conditions. To clarify the nature of this canonical variate the individual (canonical) transients for each condition are shown together on the right. It is immediately obvious that the principal difference among the three conditions was the effect of moving (i.e., pronounced response to the fixed and random conditions with little contribution from the rest condition). More remarkable, however, is the difference between the fixed and the random conditions. The system defined by the canonical images responds to both the movement conditions with complicated and biphase transients. The fixed condition evokes a rapid phasic response that adapts quickly after about 12 s to evidence an undershoot peaking at 20 s. Conversely the random condition (requiring sustained attention) does not activate until 10 s or so after the onset of the new task. It also shows a biphase response that lags the fixed response by about 10 s. A key observation here is that the average activity integrated over the entire epoch is about the same for all three conditions and yet all three conditions evoke profoundly different transient hemodynamic responses.

Figure 4 shows the canonical variates superimposed for each epoch. The general features described above are evident and the reproducibility over time is clear. To demonstrate that this result is not a peculiarity of this subject we repeated exactly the same experiment and analysis with a second (female) subject. The first canonical variate is shown in Fig. 5. The same generic features are observed, namely, an early biphase response to the fixed condition and a more protracted biphase response to the random condition.

**DISCUSSION**

In this paper we have presented a general multivariate analysis of evoked responses and their spatiotemporal dynamics as measured with fast fMRI. This analysis uses standard multivariate statistics and the general linear model to make inferences about effects of interest and a canonical variates analysis to describe the important features of these effects. We do not propose that this analysis is a particularly powerful way to make inferences about functional specialization (or more generally functional anatomy) because it only provides statistical inferences about the distributed (whole brain) responses evoked by tasks and conditions. However, its application to characterizing the temporal aspects of evoked hemodynamic responses yields some compelling insights about transient responses to changes in cognitive or sensorimotor pro-
cessing. The key observation is that these responses can be biphasic and show profound differences in their form depending on the extant task or condition. Furthermore these differences can be seen in the absence of changes in mean signal. This has potentially severe implications for current analyses that try to find a mean difference in task-dependent signal or assume some fixed form for the evoked hemodynamic transient.

**Relationship to Conventional Analyses of fMRI Time-Series**

This multivariate approach differs fundamentally from statistical parametric mapping and related approaches because the concept of a separate voxel or region of interest ceases to have meaning. One observation represents a whole scan. Any statistical inference that ensues following a multivariate analysis is about the whole volume, not any component of it. This precludes statistical inferences about regional effects that are made without reference to changes elsewhere in the brain. This fundamental difference ensures that univariate and multivariate approaches are likely to be used in their distinct domains and should be regarded as complementary approaches. Statistical parametric mapping (e.g., $t$ maps reflecting correlations with a reference vector) allow for inferences that a small brain region is implicated in a some specified way. Multivariate tests of the sort described in this paper allow the inference that the entire brain volume imaged was engaged by the activations employed. The nature of this response is then characterized using a post hoc analysis, in this case CVA.

The CVA component proposed in this paper is conceptually similar to eigenimage analysis. Unlike eigenimage analysis CVA can be considered a true “statistical” procedure because error terms are explicitly modeled and there exist distributional approximations for the canonical values. Canonical images can be thought of as de-noised eigenimages that are informed by (and attempt to discount) error.

**CONCLUSION**

We have developed a multivariate approach to fMRI time-series using established techniques in statistics and using proven ideas in the field of evoked potentials. This aim of this analysis was to assess the spatial and temporal characteristics of evoked responses in a way that makes full use of the high spatial and temporal resolution now afforded by fMRI. The finding that the hemodynamic signal can adapt in a task-specific fashion over several seconds should not, of course, be surprising. Adaptation (and the behavioral counterpart—habituation) occurs in nearly every neural system over nearly every time-scale. Recognizing that this general principle can also apply to hemodynamic changes measured over several tens of seconds should enable a richer and more valid statistical modeling of the differences in response to different tasks. This is the subject of a companion paper that describes a univariate approach to testing for these differential responses.

**ACKNOWLEDGMENTS**

The authors are supported by the Wellcome Trust. We thank colleagues for help and support in developing these ideas, in particular our colleagues at the Institute of Psychiatry, London, UK, for essential help in doing the experiments.

**REFERENCES**


