• Human Brain Mapping 4:140–151(1996) •

A Multivariate Analysis of PET Activation Studies

Karl J. Friston, Jean-Baptiste Poline, Andrew P. Holmes, Chris D. Frith, and Richard S.J. Frackowiak

Wellcome Department of Cognitive Neurology, The National Hospital, Queen Square WC1N 3BG, UK

Abstract: In this paper we present a general multivariate approach to the analysis of functional imaging studies. This analysis uses standard multivariate techniques to make statistical inferences about activation effects and to describe the important features of these effects. More specifically, the proposed analysis uses multivariate analysis of covariance (ManCova) with Wilk's lambda to test for specific effects of interest (e.g., differences among activation conditions), and canonical variates analysis (CVA) to characterize differential responses in terms of distributed brain systems. The data are subject to ManCova after transformation using their principal components or eigenimages. After significance of the activation effect has been assessed, underlying changes are described in terms of canonical images. Canonical images are like eigenimages but take explicit account of the effects of error or noise. The generality of this approach is assured by the general linear model used in the ManCova. The design and inferences sought are embodied in the design matrix and can, in principle, accommodate most parametric statistical analyses. This multivariate analysis may provide a statistical approach to PET activation studies that 1) complements univariate approaches like statistical parametric mapping, and 2) may facilitate the extension of existing multivariate techniques, like the scaled subprofile model and eigenimage analysis, to include hypothesis testing and statistical inference. © 1996 Wiley-Liss, Inc.

Key words: multivariate analysis of covariance, canonical variates analysis, principal component analysis, PET, functional anatomy, verbal fluency

INTRODUCTION

The aim of this paper is to describe how a standard multivariate analysis can be applied to functional images in a voxel-based fashion. Specifically, we introduce multivariate analysis of covariance (ManCova) and canonical variates analysis (CVA) to characterize activation effects and address the special issues that ensue. The proposed approach characterizes the brain's response in terms of functionally connected and distributed systems. This characterization is usually associated with eigenimage analysis using singular value decomposition or principal component analysis [see Friston et al., 1993, 1994; Moeller et al., 1987, for a discussion of the conceptually related scaled subprofile model (SSM)]. Eigenimages figure in the current analysis in the following way. One problematic issue, in the multivariate analysis of functional imaging data, is that the number of samples (i.e., scans) is usually very small in relation to the number of components (i.e., voxels) of the observations. This issue is resolved by analyzing the data, not in terms of voxels, but in terms of eigenimages, where the number of eigenimages is much smaller than the number of voxels. The importance of the analysis presented in this paper is fourfold. 1) Unlike previous multivariate approaches it provides for statistical inferences (i.e., a P value) about the significance of the brain's response in terms of some hypothesis. 2) The approach implicitly takes

Received for publication December 28, 1995; accepted March 8, 1996.

Address reprint requests to Karl J. Friston, Wellcome Department of Cognitive Neurology, The National Hospital, Queen Square WC1N 3BG, UK.

© 1996 Wiley-Liss, Inc.

account of spatial correlations in the data without making any assumptions. 3) The canonical variates analysis produces eigenimages (canonical images) that capture the activation effects, while suppressing the effects of noise or error. 4) The theoretical basis is well-established and can be found in most introductory texts on multivariate analysis.

Functional mapping studies are usually analyzed with some form of statistical parametric mapping. Statistical parametric maps are spatially-extended statistical processes that are used to test hypotheses about regional effects. Statistical parametric maps (SPMs) use univariate tests at each and every voxel and are interpreted by assuming that, under the null hypothesis, they behave as smooth Gaussian fields [Friston et al., 1991; Worsley et al., 1992]. Gaussian fields are used to model the stationary spatial covariance structure typical of imaging data. Statistical inference is based on thresholding the SPM to create activation foci. This characterization of physiological responses is based on *functional specialization,* or *segregation,* as a principle of brain organization. In an attempt to assess the *functional integration* of specialized areas, an alternative approach has been suggested [Friston et al., 1993, 1994]. This approach uses the eigenimages or principal components of imaging time-series: if functional connectivity is defined as the temporal correlation between remote neurophysiological events, then eigenimages are the eigenvectors of the functional connectivity matrix. Eigenimage analysis is predicated on similar approaches in the analysis of multichannel EEG [e.g., Friedrich et al., 1991], MEG [e.g., Fuchs et al., 1992], and multiunit electrode recordings [e.g., Mayer-Kress et al., 1991]. Although powerful, in a descriptive sense, eigenimage analysis and related approaches are not generally considered as "statistical" methods that can be used to make statistical inferences; they are mathematical devices that simply identify prominent patterns of correlations or functional connectivity. It must be said, however, that large-sample, asymptotic, multivariate normal theory could be used to make some inferences about the relative contributions of each eigenimage (e.g., tests for nonsphericity) if a sufficient number of scans were available.

In what follows, we observe that multivariate analysis of covariance (ManCova), with canonical variates analysis, combines many of the attractive features of statistical parametric mapping and eigenimage analysis. Unlike statistical parametric mapping, ManCova is multivariate. In other words, it considers one observation as comprising all the voxels in a single scan. The importance of this multivariate approach is that the effects due to activations, confounding effects, and error effects are assessed both in terms of the effects at each voxel *and interactions among voxels.* This means one does not have to assume anything about spatial correlations (e.g., stationarity with Gaussian field models) when assessing the significance of the activation effect. Unlike statistical parametric mapping, these correlations are explicitly included in the analysis. The price one pays for adopting a multivariate approach is that one cannot make statistical inferences about regional changes (cf. statistical parametric mapping). This is because the inference pertains to all the components (voxels) of the multivariate variable (not a particular voxel or set of voxels).

In general, multivariate analyses are implemented in two steps. First, the significance of the hypothesized effect is assessed in terms of a P value, and secondly (if justified), the exact nature of the effect is determined. The analysis here conforms to this two-stage procedure: having assessed the brain's response to be significant using ManCova, the nature of this response remains to be characterized. We propose that canonical variate analysis (CVA) is an appropriate way to do this. The canonical images obtained with CVA are similar to eigenimages but are based on both the activation and the error effects. 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 is to think of them as eigenimages that reflect functional connectivity due to activations, while discounting spurious correlations due to error.

The paper is divided into two sections. The first section deals with the theory of ManCova and CVA. It presents the operational equations behind the multivariate general linear model and statistical inferences about activation effects based on Wilk's lambda, and it characterizes the nature of these effects using CVA. CVA is then discussed in relation to eigenimage analysis as previously implemented in functional imaging [Friston et al., 1993]. The second section is an illustrative application to a standard PET activation study of word generation in normal subjects. The data are used to compare results with those obtained using eigenimage analysis. 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], but see also Mardia et al. [1979].

THEORETICAL BACKGROUND

Dimension reduction and eigenimages

The first step in multivariate analysis is to ensure that the dimensionality (number of components or voxels) of the data is smaller than the number of observations. Clearly for images 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 scan-dependent expression of eigenimages or spatial modes as a reduced set of components for each multivariate observation (scan). The eigenimages and their associated expression, or eigenvectors, in time can be calculated in a number of ways. We use the standard eigenvalue solution in this paper (an alternative would be singular value decomposition):

$$[\epsilon\ \lambda] = \mathrm{eig}(X^* \cdot X^{*T})$$

where:

$$(X^* \cdot X^{*T}) \cdot \epsilon = \epsilon \cdot \lambda$$

$$U = X^{*T} \cdot \epsilon \cdot \lambda^{-1/2}$$

and

$$X = \epsilon \cdot \lambda^{1/2} \qquad (1)$$

Here X^* is a large matrix of corrected voxel values with one column for each voxel and one row for each scan. "Corrected" implies mean correction and the removal of any confounds using linear regression. λ is a diagonal matrix of eigenvalues, and ϵ is a unitary orthonormal matrix of eigenvectors over time. The eigenimages or spatial modes constitute the columns of U, another unitary orthonormal matrix, and their expression over scans corresponds to the columns of the matrix X. X has one column for each eigenimage and one row for each scan. In our work we use only the J columns of X and U associated with eigenvalues greater than unity (after normalizing each eigenvalue by the average eigenvalue). We present the derivation of the eigenimages in this rather clumsy way because computationally it is much easier to compute the eigenvectors of $X^*.X^{*T}$ than it is for $X^{*T}.X^*$, the latter being an exceedingly large matrix. The eigenimage decompositions based on these two products are identical. This can be seen by rearranging Eq (1) to give:

$$(X^{*T} \cdot X^*) \cdot U = U \cdot \lambda$$

Intuitively, X can be thought of as the original data X^* "looked at" from a different direction or, more formally, rotated into the subspace of the J largest eigenimages. The elements of X are x_{ij} the activity of the jth eigenimage in scan i.

General linear model and design matrix

In matrix notation, the general linear model can be written as:

$$X = G\beta + e \qquad (2)$$

The general linear model assumes the errors e are independent and identically distributed with the normal distribution $[N(0, \sigma_j^2)]$. The matrix G is called the design matrix. The design matrix has one column for every effect (factor or covariate) in the model. β is the parameter matrix with one column vector β_j of parameters for each mode. The elements of G are *explanatory* variables relating to the conditions under which the observation (e.g., scan) was made. These coefficients can be 1) covariates (e.g., global cerebral blood flow, time, plasma prolactin level, etc.), in which case Eq. (2) is a familiar multivariate regression model, or (2) indicator-type variables, taking integer values to indicate the level of a factor (e.g., condition, subject, drug, etc.) under which the response variable was measured [Chatfield and Collins, 1980]. In this case the model would correspond to a Manova. There is no mathematical distinction between covariates and indicator-type variables, and if both are present the ensuing analysis would be called a ManCova.

Least squares estimates of β, say b, are given by

$$b = (G^TG)^{-1}G^TX$$

where

$$E\{b_j\} = \beta_j \text{ and } \mathrm{Var}\{b_j\} = \sigma_j^2(G^TG)^{-1} \qquad (3)$$

$\mathrm{Var}\{b_j\}$ is the variance-covariance matrix for the parameter estimates corresponding to the jth mode. These simple equations can be used to implement a vast range of statistical analyses. The design matrix can contain both covariates and indicator variables reflecting the experimental design. Each column of G

has an associated unknown parameter in the vectors β_j. Some of these parameters will be of interest. The remaining parameters will be of no interest. This distinction suggests that G (and β) can be split into two partitions G = [H D] and similarly $\beta = [\alpha^T \gamma^T]^T$ with estimators $b = [a^T g^T]^T$. Here, effects of interest are denoted by H and confounding effects of no interest by D. Eq. (2) can be expanded:

$$X = H \cdot \alpha + D\gamma + e \quad (4)$$

where H represents a matrix of 0s or 1s depending on the level or presence of some interesting condition or treatment effect (e.g., the presence of particular cognitive component), or the columns of H might contain covariates of interest that could explain the observed variance in X (e.g., dose of apomorphine or "time on target"). D corresponds to a matrix of indicator variables denoting effects that are not of any interest (e.g., of being a particular subject or block effect), or covariates of no interest (i.e., "nuisance variables" such as global activity or confounding time effects). To make this general formulation clear, consider the model for an unpaired t-test. In this instance the elements of the column vector G are −1 for all rCBF measurements in one group and 1 for the other group. A simple regression of reaction time on rCBF would be implemented by making G a column vector containing the reaction time data. The randomized block design ANCOVA implemented by the SPM95 software corresponds to $G = [H\ D_s\ D_c]$ where H specifies the activation condition, the D_s account for subject (block) effects, and D_c is a column vector of confounding global CBF covariates. The point to be made here is that nearly every conventional statistical design can be modelled as a special case of Eq. (4).

Statistical inference

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

$$R = R(\Omega) = (X - G \cdot b)^T(X - G \cdot b) \quad (5)$$

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

$$T = (H \cdot a)^T \cdot (H \cdot a). \quad (6)$$

The error sum of squares and products under the null hypothesis $R(\Omega_0)$, i.e., after discounting the effects of interest (H), is given by:

$$R(\Omega_0) = (X - D \cdot g)^T \cdot (X - D \cdot g). \quad (7)$$

Clearly if D does not exist this simply reduces to the sum of squares and products of the response variable ($X^TX = \lambda$). The significance can now be tested with:

$$\Lambda = |R(\Omega)| / |R(\Omega_0)| \quad (8)$$

where Λ is Wilk's statistic (known as Wilk's lambda). A special case of this test is Hotelling's T^2 test and applies when H simply compares one condition with another [see Chatfield and Collins, 1980]. Under the null hypothesis, after transformation Λ has a χ^2 distribution with degrees of freedom J.h. The transformation is given by:

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

where r is the degrees of freedom associated with the error terms and is the number of scans (I) minus the number of effects modelled = I − rank(G). J is the number of eigenimages or modes in the J-variate response variable X, and h is the degrees of freedom associated with the effects of interest = rank(H).

The potential weaknesses here include the facts that 1) the χ^2 distribution is an approximation, and 2) even if the Gaussian assumptions of Eq. (2) hold for the error terms, any non-Gaussian components in the response variable that are not modelled in the design matrix may violate the distributional assumptions. There is no particular reason that PET data should be more susceptible to these weaknesses than any other data, but we mention them for completeness.

Characterizing the effect

Having established that the effects of interest are significant (e.g., differences among two or more activation conditions), the final step is to characterize these effects in terms of their spatial topography. This characterization uses canonical variates analysis (CVA). The objective is to find the linear combination (compound or contrast) of the components of X, in this case

the spatial modes or eigenimages, that best express the activation effects when compared to error effects. More exactly, we want to find c_1 such that the variance ratio:

$$(c_1^T \cdot H \cdot c_1)/(c_1^T \cdot R \cdot c_1)$$

is maximized [Chatfield and Collins, 1980]. Let $Z_1 = X.c_1$ where Z_1 is the first canonical variate and c_1 is a canonical image (defined in the space of the spatial modes) that maximizes this ratio. c_2 is the second canonical image that maximizes the ratio subject to the constraints that 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:

$$T \cdot c = R \cdot c \cdot \theta \qquad (10)$$

where Θ is a diagonal matrix of eigenvalues. Voxel-space canonical images C are obtained by rotating the canonical image in the columns of c back into voxel-space with the original eigenimages U:

$$C = U \cdot c \qquad (11)$$

The columns of C now contain the voxel values of the canonical images. The kth column of C (the kth canonical image) has an associated canonical value equal to kth leading diagonal element of Θ times r/h. Note that the "activation" effect is a multivariate one, with J components or canonical images. Normally only a few of these components have large canonical values, and only these need be reported. An idea of whether a particular canonical image is important can be assessed with its canonical value. As noted above, the canonical value corresponds to a variance ratio and can be compared to $F_\alpha(h, r)$. If the canonical value exceeds a critical threshold [e.g., $F_{0.05}(h, r)$] it might be taken seriously. However F(h, r). is not a distributional approximation for the canonical values (these values have been chosen to maximize the variance ratio). Statistical inference is based on Wilk's lambda and pertains to all the canonical images. There are procedures based on distributional approximations of Θ that do allow inferences about the dimensionality of the response (number of canonical images). We refer the interested reader to Chatfield and Collins [1980] for further details.

Relationship to eigenimage analysis

In Friston et al. [1993] we proposed that the eigenvectors of the covariance matrix based on the adjusted condition means following an ANCOVA were a useful characterization of the functional interactions observed in an activation study. "Adjusted condition means" refer to the means for each condition after confounds, such as subject effects, have been removed. When the design matrix partition H models only these conditions, the adjusted condition means are the same as the parameter estimates in the matrix a above (expressed in terms of the spatial modes U). The eigenvectors of the covariance of the adjusted condition means correspond to the eigenvectors of $a^T.a$. $a^T.a$ is directly proportional to $T = a^T.H^T.H.a$ ($H^T.H$ is proportional to the identity matrix by orthogonality of the experimental design). Therefore, the eigenimages in Friston et al. [1993] correspond to the eigenvectors of T. These have an interesting relationship to the canonical images: On rearranging Eq. (10), i.e.:

$$R^{-1} \cdot T \cdot c = c \cdot \theta$$

we note that the canonical images are the eigenvectors of $R^{-1}.T$. In other words, an eigenimage analysis of an activation study (as proposed in Friston et al. [1993]) returns the eigenvectors that express the most variance due to the effects of interest – eig(T). A canonical image, on the other hand, expresses the greatest amount of variance due to the effects of interest *relative to error* – eig($R^{-1}.T$). In this sense, a CVA can be considered an eigenimage analysis that is "informed" by the estimates of the error effects.

AN ILLUSTRATIVE APPLICATION

In this section, we consider an application of the above theory to a word generation study in normal subjects. We will use this illustrative example to comment further on the various implementations of the multivariate approach and the interpretation of canonical images.

The data

The data were obtained from 5 subjects scanned 12 times (every 8 min) while performing one of two verbal tasks. Scans were obtained with a CTI PET camera (model 953B, CTI, Knoxville, TN) [Spinks et al., 1992]. ^{15}O was administered intravenously as radiolabelled water infused over 2 min. Total counts per voxel during the buildup phase of radioactivity served as an estimate of regional cerebral blood flow (rCBF). Subjects performed two tasks in alternation. One task involved repeating a letter presented aurally at one

per 2 sec (*word shadowing*). The other was a paced verbal fluency task, where subjects responded with a word that began with the letter presented (*intrinsic word generation*). To facilitate intersubject pooling, the data were realigned and stereotactically normalized [Friston et al., 1995]. Intracranial voxels were selected (using an arbitrary threshold criterion), and mean-corrected and adjusted for subject, time, and global activity using linear regression, as described in Friston et al. [1995]). These data constitute the data matrix X*.

ManCova

We assessed the significance of condition-dependent effects by treating each of the 12 scans as a different condition. Note that we do not consider the six word-generation (or word-shadowing) conditions as replications of the same condition. In other words, the first time one performs a word-generation task is a different condition from the second time, and so on. The (alternative) hypothesis adopted here states that there is a significant difference among any of the 12 conditions but does not constrain the nature of this difference to a particular form. The most important differences will emerge from the CVA. Clearly one might hope that these differences will be due to word generation, but they might not be. This hypothesis should be compared with a more constrained hypothesis that considers the conditions as six replications of word shadowing and word generation. This hypothesis is more directed and explicitly compares word shadowing with word generation. This comparison could be tested in a single subject. The point being made here is that the generality afforded by the current framework allows one to test very constrained (i.e., specific) hypotheses or rather general hypotheses about some unspecified activation effect. We choose the latter here because it places more emphasis on the canonical images as descriptions of what has actually occurred during the experiment.

The design matrix partition for effects of interest H had 12 columns representing the 12 different conditions. We designated subject effects, time, and global activity as uninteresting confounds. The partition D of the design matrix therefore had four columns for each subject (subject effects were constrained to zero, eschewing the need to incorporate the fifth subject-effect explicitly), five columns for time effects in each subject, and one column of global activities (even if these confounds have already been removed in a preprocessing step, it is important to use them again here because fitting confounds alone, and fitting them as part of a complete model, will give different

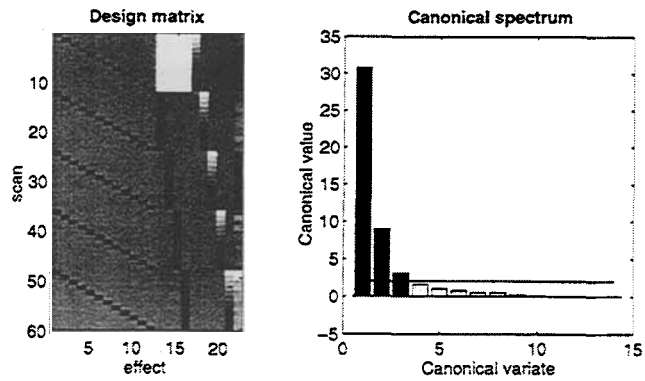


Figure 1.

Left: Design matrix used in demonstration analysis. Design matrix G models 22 effects. The 12 condition effects are at left, and 10 confounds are at right). These confounds include subject effects, time, and global activity. Matrix is displayed in image format, with each column scaled to its absolute maximum. Design matrix shows that 60 scans are ordered as 12 conditions from subject 1, followed by 12 conditions from subject 2, and so on to subject 5. **Right:** Spectrum of canonical values following a canonical variates analysis of the sums of squares and product matrices due to condition and error terms. Black bars represent canonical values that exceed $F_{0.05}(h, r)$.

estimates). The complete design matrix is seen in Figure 1 (left) and is displayed in image format. The condition effects are seen in the left part of the design matrix, and the confounds on the right.

The corrected data were reduced to 60 eigenvectors, as described in Theoretical Background. The first 14 eigenvectors had eigenvalues greater than unity and were used in the subsequent analysis. The resulting matrix X, with 60 rows (one for each scan) and 14 columns (one for each eigenimage), was subject to ManCova. The significance of the condition effects was assessed with Wilk's Lambda. According to Eq. (8), the P value for the condition or activation effects was $P = 0.02$. In other words, the probability of no differences among the 12 conditions was 0.02.

Canonical variates analysis

The condition effects were almost completely accounted for by two canonical images. In other words, two canonical values were substantially larger than $F_{0.05}(h, r) = F_{0.05}(11, 37) = 2.4$. The spectrum of canonical values is seen on the right in Figure 1. The first canonical image and its expression in each condition are shown in Figure 2. Upper panels show this system to include anterior cingulate and Broca's area, with more moderate expression in the left posterior inferotemporal regions (right). The positive components of

this canonical image (left) implicate the ventromedial prefrontal cortex and the bitemporal regions (right greater than left). One important aspect of these canonical images is their highly distributed yet structured nature, reflecting the distributed integration of many brain areas.

The canonical variate expressed in terms of mean condition effects is seen below in Figure 2. This variate is simply $a.c_1$. It is pleasing to note that the first canonical variate corresponds to the difference between word shadowing and verbal fluency. The corresponding canonical image is clearly implicated in the difference between the activation conditions (odd) and baseline (even) and is consistent with the known functional anatomy of verbal fluency.

The canonical variate expressed over scans is shown in Figure 3 and is given by $Z_1 = X.c_1 = X^*.C$. For convenience, the expression of Z_1 is shaded to highlight scans from the different subjects. Z_1 is, almost universally, high negative in word generation and high positive in word shadowing. There are substantial differences in the expression of this canonical image over time within each subject, reflecting idiosyncratic time-dependent changes in activation.

Comparison with eigenimage analysis

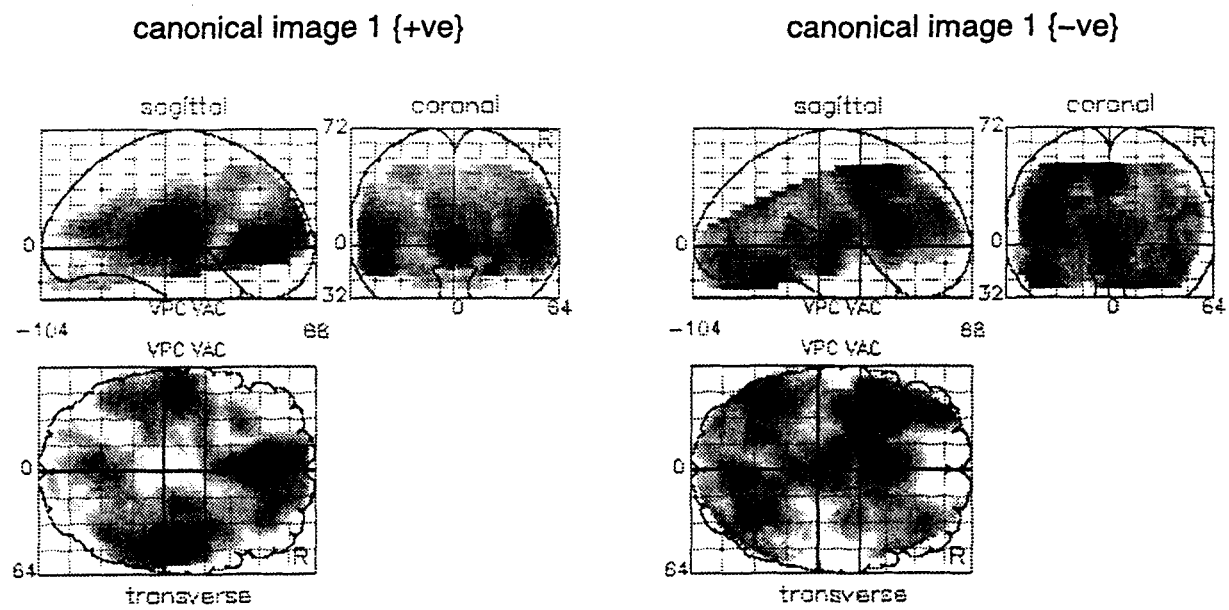
The first eigenimage of the condition sum of squares and products matrix T is shown in Figure 4 (after rotation back into voxel-space), and corresponds to the eigenimage analysis described in Friston et al. [1993]. This eigenimage reflects the main patterns of correlations evoked by the mean condition effects and should be compared with the first canonical image in Figure 2. The eigenimage in the current figure again implicates the anterior cingulate, Broca's area, the left posterior inferotemporal regions, and the bitemporal regions. In addition, another area has appeared, i.e., the posterior cingulate. In some ways the eigenimage is more compelling than the canonical image, but the differences between these characterizations of activation effects are informative. The eigenimage is totally insensitive to the reliability or error attributable to differential activation from subject to subject, whereas the canonical image does reflect these variations. For example, the absence of the posterior cingulate in the canonical image and its relative prominence in the eigenimage suggest that this region is implicated in some subjects but not in others. The subjects that engage the posterior cingulate must do so to some considerable degree because the average effects (represented by the eigenimage) are quite substantial. Conversely, the medial prefrontal cortical deactivations are a much more generic feature of activation effects than would have been inferred on the basis of the eigenimage alone. These observations beg the question, "Which is the best characterization of functional anatomy?" Obviously there is no simple answer, but the question speaks to an important point: the canonical image characterizes the response *relative to error,* by partitioning the observed variance (in the J larger spatial modes) into effects we are interested in and a residual variation about these effects (error). This partitioning is determined by the experimental design and the inferences that are sought. The eigenimage does not embody any concept of error and is not constrained by any hypothesis.

Error terms

Figure 5 shows the first eigenimage of the sum of squares and products due to error R. The topography of this eigenimage is highly structured and appears to reflect a reciprocal relationship between the cortical envelope and the brain's interior. It is also interesting to note that the striate cortex and right temporoinsular regions are heavily implicated, suggesting that variability in these regions is high and strongly coupled. With the current model the error terms can be thought of as interactions between condition and subject effects, in other words, as the subject-to-subject differences in condition-dependent responses. This being the case, one might conjecture that the error terms will embody those condition effects that are expressed more in some subjects than in others. The expression of this error eigenimage concurs with this conjecture. The lower panel shows the expression of the first eigenimage of R in the same format as in Figure 3. It can be seen that subjects 3 and 4 express this mode during word generation, whereas the remaining subjects do not. The expression in subject 5 suggests a highly nonlinear time effect. This sort of effect illustrates the potential for non-Gaussian components in the error terms and, if very prevalent, should be noted as a violation of the distributional assumptions implicit in the general linear model. Alternatively they could, of course, be appropriately modelled in the design matrix.

DISCUSSION

In this paper we have presented a general multivariate approach to the analysis of functional imaging studies. This analysis uses standard multivariate techniques to make statistical inferences about activation effects and to describe the important features of these



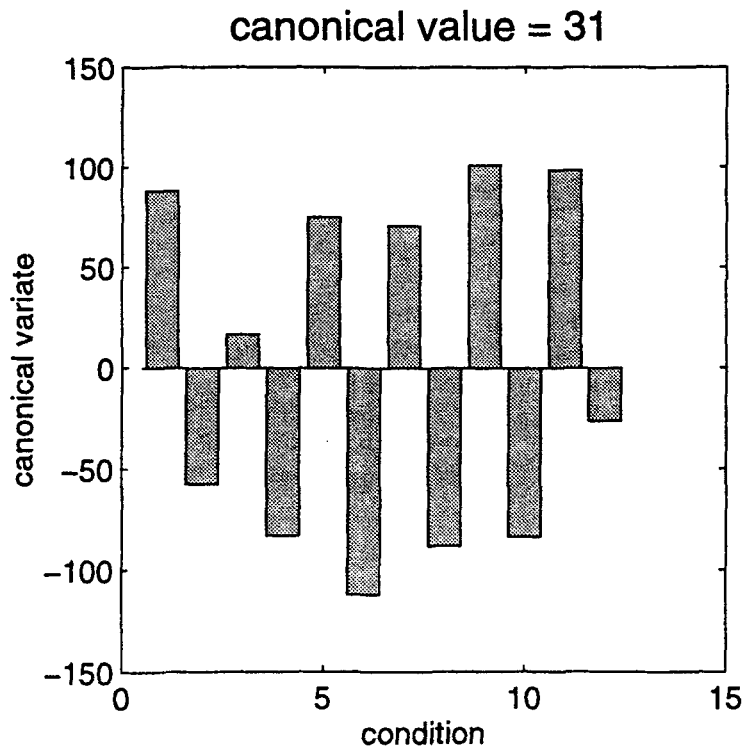


Figure 2.

Top: First canonical image, displayed as maximum intensity projections of positive and negative components. Display format is standard and provides three views of the brain from front, back, and right side. Grayscale is arbitrary, and the space conforms to that described in Talairach and Tournoux [1988]. **Bottom:** Expression of first canonical image (i.e., the canonical variate) averaged over conditions. Odd conditions correspond to word shadowing, and even conditions correspond to word generation. This canonical variate is clearly sensitive to the differences evoked by these two tasks.

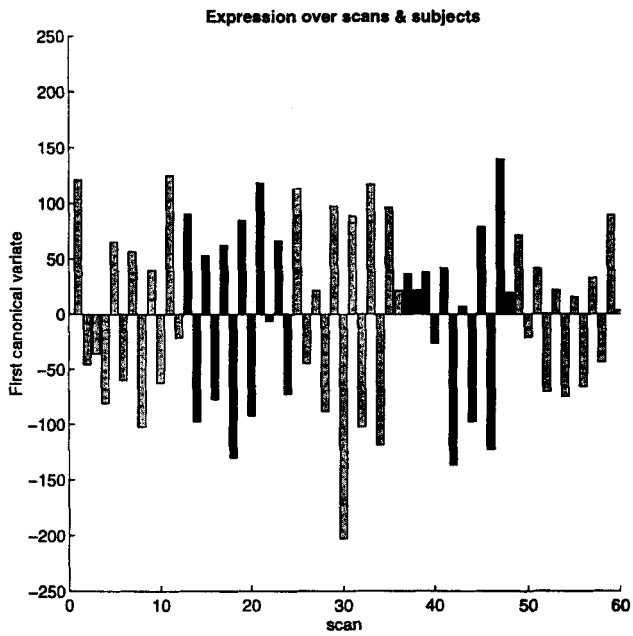


Figure 3.
Canonical variate Z_1 expressed over all conditions, in all subjects. The scans are ordered as in Figure 2, but for each subject in turn. Scans from different subjects are alternately grey and black. It can be seen that the relationship of this canonical variate to word generation is largely preserved from subject to subject.

effects. More specifically, the proposed analysis uses multivariate analysis of covariance ManCova with Wilk's lambda to test for specific effects of interest (e.g., differences among activation conditions) and canonical variates analysis (CVA) to characterize these distributed responses. These established methods are applied to the data after they have been transformed using the underlying principal components or eigenimages. After the significance of the activation effect has been assessed, the underlying distributed changes are described in terms of canonical images. The analysis can be summarized in terms of the following stages:

- Reduce the dimensionality of the data (equal to number of voxels) by rotating the data into the eigenimage space (with a dimensionality that is less than the number of scans).
- Assess the significance of interesting (e.g., activation) effects using an appropriately configured design matrix, ManCova, and Wilk's lambda.
- Characterize these effects using a canonical variates analysis in terms of canonical vectors that best capture the activation effects, relative to error.
- Rotate the canonical vectors back into voxel-space. The expression of the resulting canonical images is given by the canonical variates.

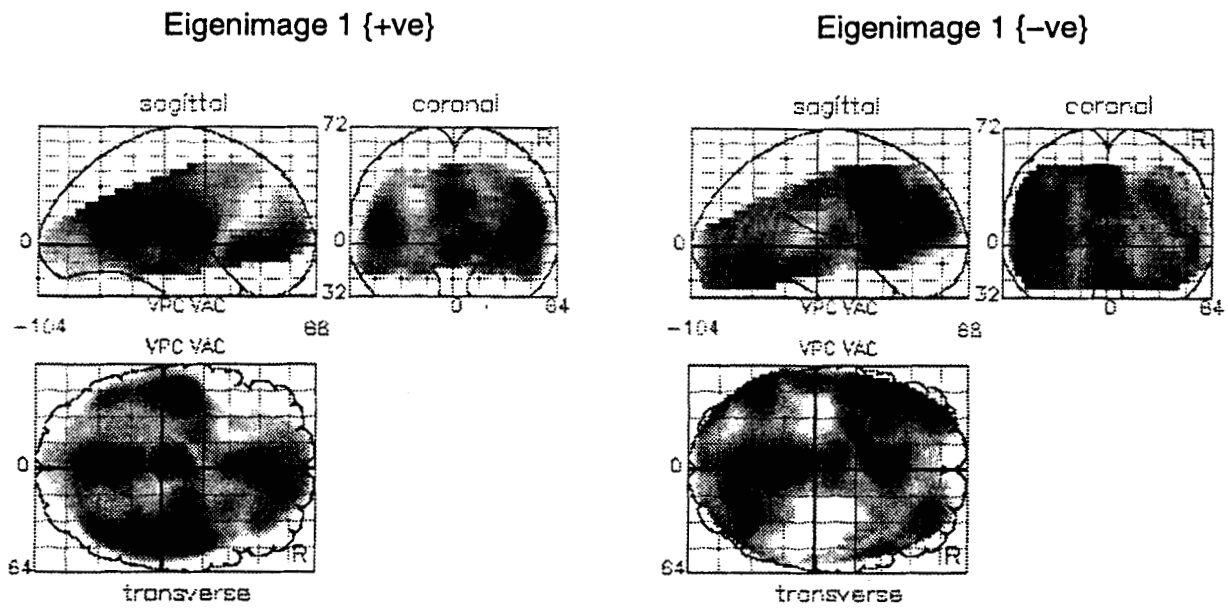
The generality of this approach is assured by the generality of the linear model used. The design and inferences sought are embodied in the design matrix and can, in principle, accommodate most parametric statistical analyses.

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 scan represents one observation (not 10^5 voxels). In this sense, the statistical inference is about the whole image volume, and 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 SPM and multivariate approaches are likely to be regarded as distinct and complementary approaches to functional imaging data.

The CVA component proposed in this paper is conceptually similar to eigenimage analysis, but can be considered a true "statistical" procedure. The reason that CVA is considered "statistical" is that the underlying mathematical model includes error terms. Canonical images can be thought of as denoised eigenimages that are informed by (and attempt to discount) error. Because canonical images are single (multivariate) objects, there are no thresholds. This may present something of a challenge to those who are used to working with thresholded SPMs and the discrete foci that ensue. Clearly, like eigenimages and subprofiles from SSM, canonical images should be reported and described in their entirety as single distributed profiles. The verbal description of a canonical image will be more anecdotal than the corresponding point-list summaries of SPM maxima. However, the canonical image itself is as valid as an SPM as a description of a significant effect (some might say more so, given that it is not subject to an arbitrary threshold).

Applications

There are many potential applications of the analysis presented in this paper. One particularly interesting application concerns the ability to test various models in a comprehensive and direct fashion. Hitherto there has been no "omnibus" test for a particular neurophysiological response or model of this response that did not rely on some assumptions about the multivariate structure of the data (e.g., Gaussian fields). Wilk's statistic could provide this test. For example, the controversy over the appropriate model for removing the confounding effects of global activity



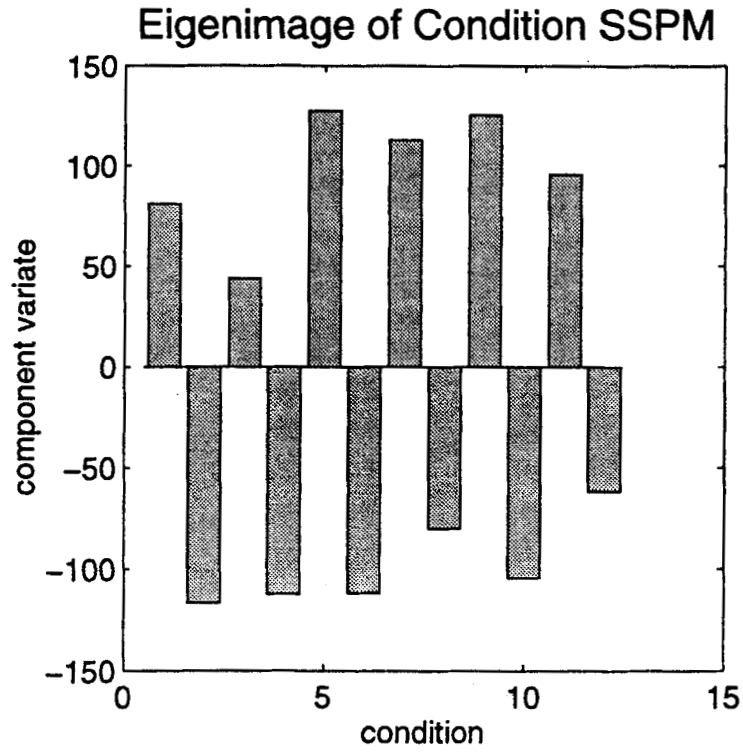
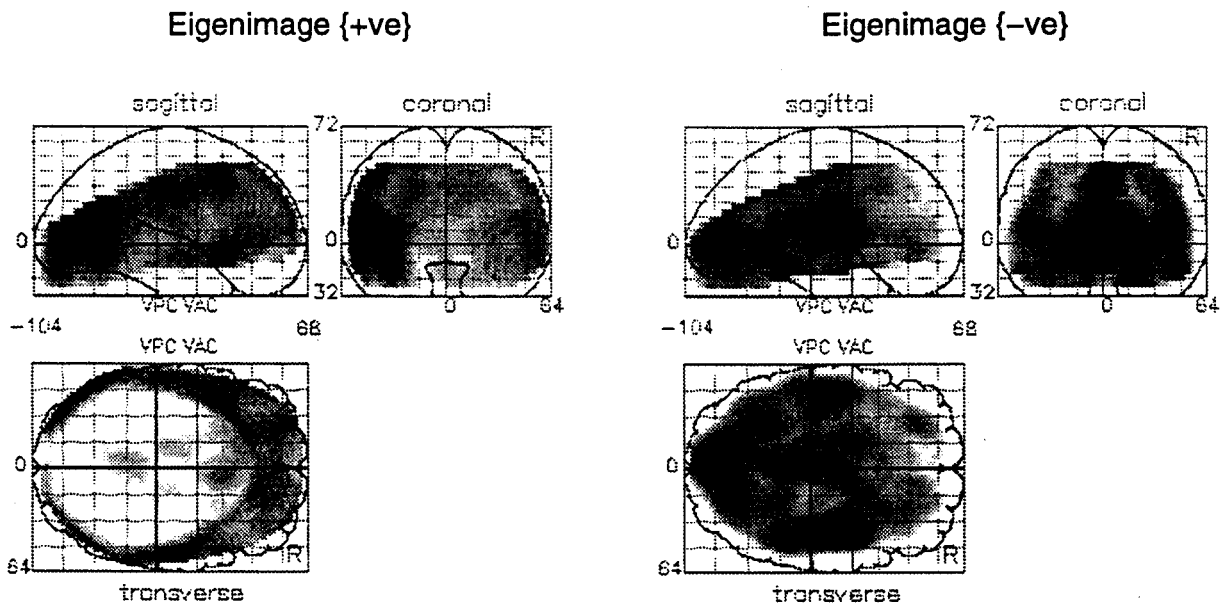


Figure 4.

First eigenimage of the sums of squares and products of the condition effect T. Display format is the same as in Figure 2. The expression of this mode is displayed below, and concurs with the expression of the first canonical image.



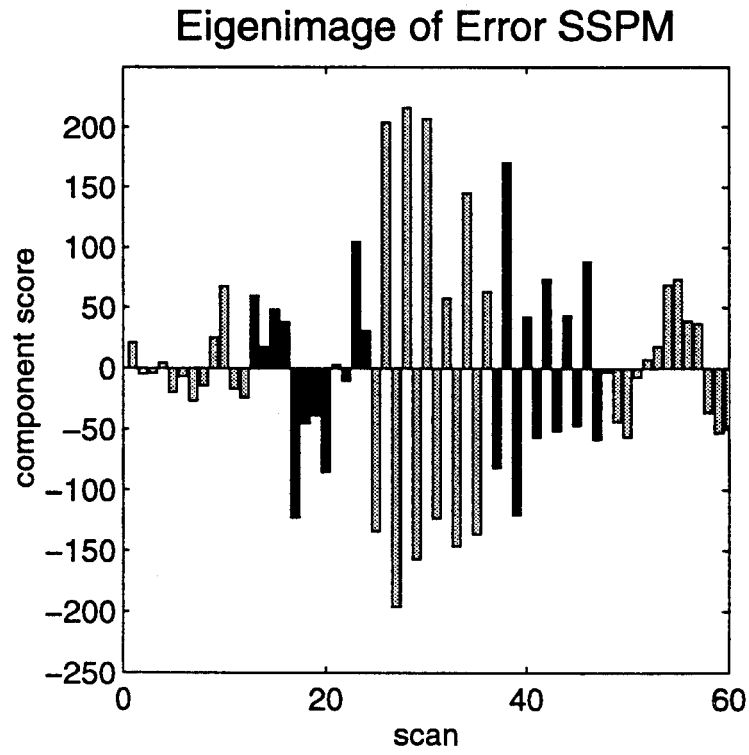


Figure 5.
First eigenimage of the sums of squares and products of the error effect R. Display format is the same as in Figure 4, but the expression of this eigenimage (below) is shown over all 60 scans, using the format of Figure 3.

on regional effects has been dogged by the lack of any compelling comparative assessment of different models. Wilk's statistic could, in principle, be used to resolve this issue by explicitly testing hierarchies of models (a succession of extra effects modelled in the design matrix).

An attractive neuroscience application of the multivariate approach considered here pertains to the significance of interaction terms in the design matrix. Cognitive subtraction is based on the assumption that extra components of a task can be inserted without affecting the preexisting components. One of the main concerns with cognitive subtraction and additive factors logic can be reduced to the relationship between neural dynamics and cognitive processes. For example, even if a cognitive component can be added without interacting with preexisting components, the brain's implementation is almost certainly going to show profound interactions. This follows from the observation that neural dynamics are nonlinear. Indeed, nearly all theoretical and computational neurobiology is based on this. In order to verify the assumptions behind cognitive subtraction, one needs to demonstrate that these interactions can be ignored when modelling the brain's response. This can be effected simply and rigorously using Wilk's statistic to show that the interaction terms in the design matrix are not significant (here one would treat the interaction terms as effects of interest and the remaining effects as of no interest). Of course, if the interactions were significant this would lead to a richer understanding of functional anatomy and to a basis for more sophisticated experimental designs.

CONCLUSIONS

We have presented a simple multivariate analytic approach to functional imaging data. This multivariate analysis may provide a statistical approach to PET activation studies that 1) complements univariate approaches like statistical parametric mapping, and 2) may facilitate the extension of existing multivariate techniques to include hypothesis-testing and statistical inference.

ACKNOWLEDGMENTS

K.J.F., A.P.H., and R.S.J.F. were funded by the Wellcome Trust. We thank all our colleagues for help and support in developing these ideas, in particular Stephen Strother. We are indebted to the anonymous reviewers for some of the key observations above.

REFERENCES

Chatfield C, Collins AJ (1980): Introduction to Multivariate Analysis. London: Chapman and Hall.

Friedrich R, Fuchs A, Haken H (1991): Modelling of spatio-temporal EEG patterns. In: Dvorak I, Holden AV (eds): Mathematical Approaches to Brain Functioning Diagnostics. New York: Manchester University Press.

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1991): Comparing functional (PET) images: The assessment of significant change. J Cereb Blood Flow Metab 11:690–699.

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1993): Functional connectivity: The principal component analysis of large (PET) data sets. J Cereb Blood Flow Metab 13:5–14.

Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC (1994): Assessing the significance of focal activations using their spatial extent. Hum Brain Mapping 1:214–220.

Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ (1995): Statistical parametric maps in functional imaging: A general linear approach. Hum Brain Mapping 2:189–210.

Friston KJ, Ashburner J, Poline JB, Frith CD, Heather JD, Frackowiak RSJ (1995): Spatial realignment and normalization of images. Hum Brain Mapping 2:165–189.

Fuchs A, Kelso JAS, Haken H (1992): Phase transitions in the human brain: Spatial mode dynamics. Int J Bifurc Chaos 2:917–939.

Mardia KV, Kent JT, Bibby JM (1979): Multivariate Analysis. London: Academic Press.

Mayer-Kress G, Barczys C, Freeman W (1991): Attractor reconstruction from event-related multi-electrode EEG data. In: Dvorak I, Holden AV (eds): Mathematical Approaches to Brain Functioning Diagnostics. New York: Manchester University Press. pp 315–336.

Moeller JR, Strother SC, Sidtis JJ, Rottenberg DA (1987): Scaled subprofile model: A statistical approach to the analysis of functional patterns in positron emission tomographic data. J Cereb Blood Flow Metab 7:649–658.

Spinks TJ, Jones T, Bailey DL, Townsend DW, Grootnook S, Bloomfield PM, Gilardi MC, Casey ME, Sipe B, Reed J (1992): Physical performance of a positron tomograph for brain imaging with retractable septa. Phys Med Biol 37:1637–1655.

Talairach J, Tournoux P (1988): A Co-Planar Stereotaxic Atlas of a Human Brain. Stuttgart: Thieme.

Worsley KJ, Evans AC, Marrett S, Neelin P (1992): A three-dimensional statistical analysis for rCBF activation studies in human brain. J Cereb Blood Flow Metab 12:900–918.