ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 37 (2007) 855–865

Local Linear Discriminant Analysis (LLDA) for group and region of interest (ROI)-based fMRI analysis

Martin J. McKeown,[a,b,c,*] Junning Li,[d] Xuemei Huang,[e] Mechelle M. Lewis,[e] Seungshin Rhee,[f] K.N. Young Truong,[f] and Z. Jane Wang[c,d]

[a]*Pacific Parkinson's Research Centre, University of British Columbia, Vancouver, Canada*
[b]*Department of Medicine (Neurology), University of British Columbia, Vancouver, Canada*
[c]*Brain Research Centre, University of British Columbia, Vancouver, Canada*
[d]*Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada*
[e]*Department of Neurology, School of Medicine, University of North Carolina, Chapel Hill, NC 27599, USA*
[f]*Department of Biostatistics, School of Public Health, University of North Carolina, Chapel Hill, NC 27599, USA*

Received 17 October 2006; revised 18 March 2007; accepted 19 April 2007
Available online 26 May 2007

A post-processing method for group discriminant analysis of fMRI is proposed. It assumes that the fMRI data have been pre-processed and analyzed so that each voxel is given a statistic specifying task-related activation(s), and that individually specific regions of interest (ROIs) have been drawn for each subject. The method then utilizes Local Linear Discriminant Analysis (LLDA) to jointly optimize the individually-specific and group linear combinations of ROIs that maximally discriminates between groups (or between tasks, if using the same subjects). LLDA tries to linearly transform each subject's voxel-based activation statistics within ROIs to a common vector space of ROI combinations, enabling the relative similarity of different subjects' activation to be assessed. We applied the method to data recorded from 10 normal subjects during a motor task expected to activate both cortical and subcortical structures. The proposed method detected activation in multiple cortical and subcortical structures that were not present when the data were analyzed by warping the data to a common space. We suggest that the method be applied to group fMRI data when warping to a common space may be ill-advised, such as examining activation in small subcortical structures susceptible to mis-registration, or examining older or neurological patient populations.

Keywords: Discriminant analysis; fMRI; Group analysis; Regions of interest

Introduction

Group analysis in fMRI is typically done in several consecutive steps. First, fMRI data are corrected for motion, despite the fact that most methods cannot easily distinguish changes in fMRI signal from that induced by motion (Liao et al., 2005, 2006). Data are then spatially transformed to a common space such as the atlas by Talaraich (Talairach and Tournoux, 1988) or the probabilistic space suggested by the Montreal Neurological Institute (Collins et al., 1998) to minimize intersubject differences. However, because of the variability in human brain anatomy, the inter-subject registration is typically imperfect, so spatial low-pass filtering ("smoothing") is performed to de-emphasize anatomical differences (Friston, 1996). Once data have been motion corrected, warped to a common space, and spatially smoothed, the task-related activation of a voxel of a subject k is estimated with linear regression techniques:

$$Y_k = \mathbf{X}_k\beta_k + \varepsilon_k, \text{ and } \mathrm{Cov}(\varepsilon_k) = \sigma_k^2\mathbf{V}_k \quad (1)$$

where Y_k is the $T_k \times 1$ time course of the voxel, $\mathbf{X}_k$ is the $T_k \times D$ design matrix containing the hypothesized activation (often incorporating estimates of the hemodynamic response function) as well as other covariates, ε_k is the $T_k \times 1$ vector of residuals, σ_k^2 is the homogeneous variance of the residuals, and $\mathbf{V}_k$ is the correlation matrix. The subscript k indicates that all the variables are related to subject k.

As fMRI data are typically not temporally white, data are often pre-whitened using a whitening matrix $\mathbf{W}_k$ such that:

$$\mathbf{W}_k\mathbf{V}_k\mathbf{W}_k^T = I \quad (2)$$

(for an excellent summary the reader is referred to: Mumford and Nichols, 2006). If each term in Eq. (1) is pre-multiplied by $\mathbf{W}_k$, we have:

$$Y_k^* = \mathbf{X}_k^*\beta_k + \varepsilon_k^* \quad (3)$$

* Corresponding author. Pacific Parkinson's Research Centre, University of British Columbia (UBC) M31, Purdy Pavilion, University Hospital, UBC Site, 2221 Wesbrook Mall, Vancouver, British Columbia, Canada V6T 2B5. Fax: +1 604 822 7866.
E-mail address: mmckeown@interchange.ubc.ca (M.J. McKeown).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2007.04.072

where the superscript *denotes the whitened quantities. The whitening matrix $\mathbf{W}_k$ is estimated by the residuals ε_k and $\mathbf{V}_k$ as:

$$\mathbf{W}_k = \mathbf{V}_k^{-\frac{1}{2}}. \tag{4}$$

The regression estimates of Eq. (3) can then be estimated by Ordinary Least Squares (OLS) to give the Generalized Least Squares estimate of Eq. (1):

$$\hat{\beta}_k^{\mathrm{GLS}} = (\mathbf{X}_k^{*T}\mathbf{X}_k*)^{-1}\mathbf{X}_k^{*T}Y_k* \tag{5}$$

$$\mathrm{Cov}\left(\hat{\beta}_k^{\mathrm{GLS}}\right) = \sigma_k^2(\mathbf{X}_k^{*T}\mathbf{X}_k*)^{-1} \tag{6}$$

Contrasts between conditions are of most interest in an experiment, e.g., contrasting BOLD signal during performance of a given task compared to rest. In the study on a single subject, the null hypothesis is that the contrast between the least-squared estimates is zero:

$$H_0 : \boldsymbol{c}\beta_k = 0$$

where $\boldsymbol{c}$ is the contrast row vector. For example, if we are interested in the comparison between task 1 and task 2, $\boldsymbol{c}$ is [1, −1].

Group analyses are usually done using a Summary Statistics method, which is a two-staged approach; first individual models are fit to each subject as described above, and then a second level is applied to make group inferences on the $\boldsymbol{c}\beta_k$ (Mumford and Nichols, 2006). In the usual situation where one is contrasting activation across two groups, the second level is a multivariate regression equation with the design matrix encoded with group inclusion indicators (Fig. 1(a)):

$$\beta_{\mathrm{cont}} = \mathbf{X}_{\mathrm{g}}\beta_{\mathrm{g}} + \varepsilon_{\mathrm{g}} \tag{7}$$

where $\mathbf{X}_{\mathrm{g}} = \begin{pmatrix} 10 \\ \cdots \\ 10 \\ 01 \\ \cdots \\ 01 \end{pmatrix}$ is a binary $K\times 2$ matrix coded to show group inclusion (K is the number of subjects from the two groups), β_{cont} is composed of the contrasts $\boldsymbol{c}\beta_k$ for each individual as defined in the first stage, $\beta_{\mathrm{g}}=[\beta_{\mathrm{g}1}, \beta_{\mathrm{g}2}]^{\mathrm{T}}$ is mean activation of the two groups and $\varepsilon_{\mathrm{g}} \sim N(0,\delta_{\mathrm{g}}^2 V_{\mathrm{g}})$ is the residual with the variance δ_{g}^2 and the correlation matrix $\mathbf{V}_{\mathrm{g}}$ being a diagonal matrix, typically just I. Here the null hypothesis is that the group activations for a given voxel in the common spatially transformed space are not significantly different:

$$H_0 : \beta_{\mathrm{g}1} - \beta_{\mathrm{g}2} = 0.$$

A number of different implementations have been proposed to implement the above analysis in a practical way. The fMRIStat method uses Restricted Maximum Likelihood (ReML) to estimate σ_{g}^2 (Worsley et al., 2002), then smoothes the data to increase its degree of freedom and accuracy and finally tests the hypothesis with t-statistics. The SPM2 package (Friston et al., 2002a,b) estimates the $\delta^2\mathbf{V}+\delta_{\mathrm{g}}^2\mathbf{V}_{\mathrm{g}}$ term with ReML under a simplifying assumption that all the subjects share a common covariance matrix $\delta_k^2\mathbf{V}_k=\delta^2\mathbf{V}$, and then tests the hypotheses with F statistics. The FMRIB software library estimates σ_{g}^2 with the maximum *a posteriori* (MAP) criteria, then screens obviously insignificant voxels with Z-statistics and finally performs a Bayesian inference on the significance of the remaining voxels with a slower but more accurate Markov Chain Monte Carlo (MCMC) simulation (Beckmann et al., 1998).

Nevertheless, there are a number of shortcomings with the previously described methods. The above methods work on the voxel level—this assumes that after suitable spatial transformation, there is a perfect correspondence between the same voxel across subjects. While this may be mitigated somewhat by spatial smoothing, such low-pass filtering degrades the spatial resolution of the data. Activation estimates in small, subcortical structures such as the basal ganglia or thalami, which abut functionally different tissues (e.g., the internal capsule), may be particularly affected by mis-registration errors.

One way to partially circumvent the difficulties associated with spatially transforming functional maps to a common space is to manually draw anatomical regions of interest (ROIs) for each subject, and performing analyses at the ROI level—as opposed to the individual voxel level. Using standard atlases, a particular brain region (e.g., the lateral cerebellar hemisphere) is manually circumscribed on the high-resolution structural MRI scans that have been co-registered with the functional data, and the voxels within this region are analyzed. The benefit of this method is that it does not require rigid spatial transformation, preventing possible gross distortion of a particular brain area, as may occur if the anatomy of a given individual differs significantly in size and shape to the homologous area in the exemplar brain. However, drawing ROIs is labor-intensive, subject to human error, and requires the assumption that a functionally active region (the SMA for example) of a given brain will be within an anatomically standardized index (i.e., Broadman's Area 9) which is used to draw the ROI.

In addition to the possibilities of mis-registration, the previously described voxel-based methods do not explicitly model interactions between brain regions. Covarying regions are often of interest, but are not included in the group methods described above. Conceptually, group methods are done in two stages: in the first stage, individually specific regression models are fit to the data; and then the results of these models are used in a group-level analysis. Because the goal of these methods is to test a specific hypothesis, these methods may be conducted sequentially. In contrast, if the goal is to find which combination of brain regions is maximally different between tasks, it is desirable to jointly optimize the individual statistical model and the overall models simultaneously.

In individual-subject fMRI analysis, in addition to hypothesis driven methods, there is a role for data driven methods, such as Independent Component Analysis (ICA) (McKeown et al., 1998; Calhoun et al., 2003), which do not need rigorous *a priori* specification of activation patterns. In an analogous manner, there may be particular interest in discovering the combinations of brain regions (specified by ROIs) that are maximally contrasted during performance of certain tasks (Fig. 1(b)). There is therefore a need for a multivariate, discriminant analysis approach that works at the region of interest (ROI) level as opposed to the individual voxel level.

Previous work has taken individual activations (or the t-statistics associated with them) and used a multivariate discriminant approach (McKeown and Hanlon, 2004). In order to apply a discriminant approach, we first assume that some statistical analysis has been performed to assign a t-statistic, related to