NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 37 (2007) 1250 – 1259

Unsupervised analysis of fMRI data using kernel canonical correlation

David R. Hardoon,[a,*] Janaina Mourão-Miranda,[b] Michael Brammer,[b] and John Shawe-Taylor[a]

[a]*The Centre for Computational Statistics and Machine Learning, Department of Computer Science, University College London, UK*
[b]*Brain Image Analysis Unit, Centre for Neuroimaging Sciences (PO 89), Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK*

Received 22 February 2007; revised 9 June 2007; accepted 25 June 2007
Available online 3 July 2007

We introduce a new unsupervised fMRI analysis method based on kernel canonical correlation analysis which differs from the class of supervised learning methods (e.g., the support vector machine) that are increasingly being employed in fMRI data analysis. Whereas SVM associates properties of the imaging data with simple specific categorical labels (e.g., −1, 1 indicating experimental conditions 1 and 2), KCCA replaces these simple labels with a label vector for each stimulus containing details of the features of that stimulus. We have compared KCCA and SVM analyses of an fMRI data set involving responses to emotionally salient stimuli. This involved first training the algorithm (SVM, KCCA) on a subset of fMRI data and the corresponding labels/label vectors (of pleasant and unpleasant), then testing the algorithms on data withheld from the original training phase. The classification accuracies of SVM and KCCA proved to be very similar. However, the most important result arising form this study is the KCCA is able to extract some regions that SVM also identifies as the most important in task discrimination and these are located manly in the visual cortex. The results of the KCCA were achieved blind to the categorical task labels. Instead, the stimulus category is effectively derived from the vector of image features.

Keywords: Machine learning methods; Kernel canonical correlation analysis; Support vector machines; Classifiers; Functional magnetic resonance imaging data analysis

Introduction

Recently, machine learning methodologies have been increasingly used to analyse the relationship between stimulus categories and fMRI responses (Cox and Savoy, 2003; Carlson et al., 2003; Wang et al., 2003; Mitchell et al., 2004; LaConte et al., 2005; Mourao-Miranda et al., 2005, in press; Haynes and Rees, 2005; Davatzikos et al., 2005; Kriegeskorte et al., 2006). In this paper, we introduce a new unsupervised machine learning approach to fMRI analysis, in which the simple categorical description of stimulus type (e.g., type of task) is replaced by a more informative vector of stimulus features. We compare this new approach with a standard support vector machine (SVM) analysis of fMRI data using a categorical description of stimulus type.

The methodology underlying the present study originates from earlier research carried out in the domain of image annotation (Hardoon et al., 2006), where an image annotation methodology learns a direct mapping from image descriptors to keywords. Previous attempts at unsupervised fMRI analysis have been based on Kohonen self-organising maps, fuzzy clustering (Wismuller et al., 2004; Ngan and Hu, 1999) and non-parametric estimation methods of the hemodynamic response function, such as the general method described in Ciuciu et al. (2003), kernel-PCA (Thirion and Faugeras, 2003) and probabilistic ICA/PCA analysis (Beckmann and Smith, 2004). A more recent attempt has been undertaken by Faisan et al. (2005) with the application of hidden Markov event sequence models to fMRI. These Markov events are a special class of hidden Markov models (HMMs) dedicated to the modeling and analysis of event-based random processes. O'Toole et al. (2005) have reported an interesting study which showed that the discriminability of PCA basis representations of images of multiple object categories is significantly correlated with the discriminability of PCA basis representation of the fMRI volumes based on category labels.

The current study differs from previous approaches to fMRI analysis principally in that we do not apply categorical labels (e.g., −11 contrasts) to stimuli. We employ natural images rather than simple low level objects and transform each image to a vector representation summarising its main features. We then employ kernel canonical correlation analysis to associate the vector representations of image features with their corresponding fMRI image volumes. In general, canonical correlation analysis can be seen as the problem of finding basis vectors for two sets of variables such that the correlations of the projections of the variables onto corresponding basis vectors are maximised. KCCA differs from this in that it first projects the data into a higher dimensional feature space before performing CCA. CCA (Friman et al., 2001, 2003) and KCCA (Hardoon et al., 2004a) have been

* Corresponding author. Fax: +44 20 7387 1397.
E-mail address: D.Hardoon@cs.ucl.ac.uk (D.R. Hardoon).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2007.06.017

used in previous fMRI analysis, but using only conventional categorical stimulus labels. In contrast, in this work we are interested in learning the association between complex image representations and fMRI responses to characterise these associations. The fMRI data used in the following study originated from an experiment in which the responses to stimuli were designed to evoke different types of emotional responses, pleasant or unpleasant. The pleasant images consisted of women in swimsuits while the unpleasant images were a collection of images of skin diseases. Each stimulus image was represented using Scale Invariant Feature Transformation (SIFT) (Lowe, 1999) features.

We have shown that KCCA is able to extract some of the brain regions identified by supervised methods such as SVM in task discrimination (mainly in the visual cortex) and to achieve similar levels of accuracy. We discuss some of the challenges in interpreting the results given the complex input feature vectors used by KCCA in place of categorical labels.

The paper is structured as follows. Section 2 gives a review of the fMRI data acquisition as well as the experimental design and the pre-processing. These are followed by a brief description of the scale invariant feature transformation in Section 2.5. The SVM is briefly described in Section 2.6.1 while Section 2.6.2 elaborates on the KCCA methodology. Our analysis procedure is given in Section 2.7 and the results in Section 3. We conclude with a discussion in Section 4.

Materials and methods

Subjects

fMRI data were acquired from 16 right-handed healthy US college male students (aged 20–25). According to self-report, participants did not have any history of neurological or psychiatry illness. All subjects had normal vision. All subjects gave written informed consent to participate in the study after the study was explained to them. The study was performed in accordance with the local Ethics Committee of the University of North Carolina.

Data acquisition

The data for this study were collected at the Magnetic Resonance Imaging Research Center at the University of North Carolina on a 3-T Allegra Head-only MRI system (Siemens, Erlangen, Germany). The fMRI runs were acquired using a T2$*$ sequence with 43 axial slices (slice thickness, 3 mm; gap between slices, 0 mm; TR = 3 s; TE = 30 ms; FA = 80°; FOV = 192 × 192 mm; matrix, 64 × 64; voxel dimensions, 3 × 3 × 3 mm). In each run 254 functional volumes were acquired.

Experimental design

The stimuli were presented in a block fashion. There were three different active conditions: viewing unpleasant (dermatological diseases), neutral (people) and pleasant images (female models in swimsuits) and a control condition (fixation). There were 42 imagers per category. Examples of pleasant and unpleasant are given in Tables 1 and 2, respectively (we do not show natural and fixation as we do not use these instances in our work). During the experiment, there were 6 blocks of each active condition (each consisting of 7 images volumes) alternating with control blocks (fixation) of 7 images volumes. It is important to note that throughout the paper we associate pleasant with positive and unpleasant with negative.

Pre-processing

The data was pre-processed using SPM2 (Wellcome Department of Cognitive Neurology, London, UK). We used the default SPM2 pre-processing settings. All the scans were realigned to remove residual motion effects, transformed into standard space (Talairach and Tournoux, 1988) and smoothed in space using an 8-mm Gaussian filter (FWHM). The time series of each voxel was detrended using a straight-line fit linear function. In addition, we applied a mask to select voxels defining intracerebral voxels over the whole group.

Table 1
Examples of pleasant image stimulus