Human Brain Mapping 2007

Unsupervised fMRI Analysis

David R. Hardoon[1], Janaina Mourao-Miranda[2], Michael Brammer[2], John Shawe-Taylor[1]

[1]Dept. of Computer Science, University College London

[2]Brain Image Analysis Unit, Centre for Neuroimaging Sciences, Institute of Psychiatry

Recently machine learning methodology has been used increasing to analyze the relationship between stimulus categories and fMRI responses (Mitchell et al., 2004; Janaina et al., 2006; etc). Here, we introduce a new unsupervised machine learning approach to fMRI analysis approach, 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 compared this new approach with a standard Support Vector Machine (SVM) analysis of fMRI data using a categorical description of stimulus type.

The following study differs from conventional unsupervised approaches in that we make use of the stimulus characteristics. We use kernel Canonical Correlation Analysis (KCCA) to learn the correlation between the fMRI volume and the corresponding stimulus features presented at a particular time point. CCA can be seen as the problem of finding basis vectors for two sets of variables such that the correlation of the projections of the variables onto these basis vectors are mutually maximised. KCCA first projects the data into a higher dimensional feature space before performing CCA in the new feature space. To classify a new stimulus we compute the kernel between the test and the training samples and project that kernel onto the learned semantic space. The output gives a score which can be thresholded to allocate a category to each test example.
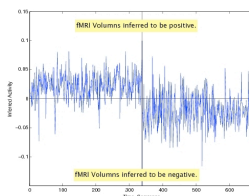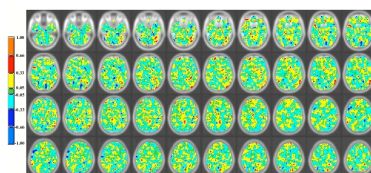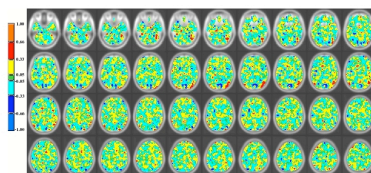
CCA (Friman et al., 2003; Friman et al., 2001) and KCCA (Hardoon et al., 2004a) have been used in previous work for fMRI analysis using only conventional categorical stimulus descriptions and did not explore the possibility of using the complex characteristics of the stimuli as the base for the feature selection in the fMRI data.

The fMRI data used is from an experiment in which we studied the responses to stimuli designed to evoke different types of emotional responses (pleasant/unpleasant). The pleasant images consisted of women in swimsuits while the unpleasant images comprised image of skin diseases. For experiment details see (Janaina et al., 2006). Each stimulus image was represented using Scale Invariant Feature Transformation (SIFT) (Lowe, 1999) features, which are used in a similar fashion to the bag-of-words (word-frequency) models for text documents.

The complete dataset consisted of 1344 fMRI volumes and their corresponding image-stimuli. From this dataset, we used 670 fMRI volumes (517,845 voxels per volume) and their corresponding image-stimuli for training and the remainder for testing. In both training/testing datasets half the stimuli were Pleasant and half Unpleasant, hence a random classifier would achieve 50% accuracy. The experiments are repeated ten

times, each with a random permutation of the training-testing set.

KCCA achieves an accuracy of 81.34%±3.29% while the supervised SVM attains an accuracy of 90.28%±0.93%. These classifications accuracy's are significantly different than of chance expectation t-test p << 0.0001. The unthresholded weights vectors for SVM are displayed in Figure 1 while those computed using KCCA are displayed in Figure 2. The similarities in the patterns of the weight vectors obtained by the two classification methods provides reassurance of the validity of the KCCA analysis, which is effectively obtaining the categorical stimulus type directly from the data. (Figure 3 gives an visual example of prediction.)

Close this window