Relevance Vector Regression Approach for Estimating Video Features from fMRI

Chia-Yueh Carlton Chu[1], Janaina Mourao-Miranda[2], Geoffrey Tan[1], John Ashburner[1]

[1]Functional Imaging Laboratory, ION, UCL, London, UK. [2]Centre for Neuroimaging Sciences, IOP, KCL, London, UK.

Abstract

The procedure involved mean removal and applying multivariate linear pattern recognition approaches (i.e. support vector regression (SVR), relevance vector regression (RVR)). The whole fMRI sequences were used, and training was done with the HRF convolved scores. Various pre processing procedures were used (ROI, masking, smoothing). A quadratic programming procedure was then carried to perform a constrained deconvolution, and re-convolution, then smoothing with the Gaussian kernel.

Introduction

The preprocessed normalized and non-normalized images provided by the competition were used. The mean of each voxel time series was subtracted. A couple of approaches were attempted in order to reduce the dimensionality of the data. The first of these involved masking out all voxels that were not grey matter. Spatial smoothing of the data was also tried. Cross-validation was used to assess the accuracy. This involved training with the first video, and testing with the second, as well as training with the second video, and testing with the first. The training and testing were done with both RVR [2] and SVR [3]. But in the end, it seems RVR provides better results. The scores provided by the web based scoring system were considers as the fitness of the approach. The best few approaches were used for the final prediction of movie 3. In the end, it was found that combining three brains together as the features actually provides better prediction for subject 1 and subject 2.

Method

Relevance Vector Regression

RVR is formulated in a Bayesian framework, and involve estimating a Restricted Maximum Likelihood (REML) solution (see e.g. [1]) of a vector of hyper-parameters **α**, **β.** These hyper-parameters are then used to estimate the best weights **w**. Given a data set of input-target pairs. $\{\mathbf{x}_n, t_n\}_{n=1}^{N}$, considering scalar-valuged target functions only. $\mathbf{x}_n$ is the voxels of the image volume. t_n is the rating convolved with the haemodynamic function.

$$\mathbf{t}_i = \sum_{n=1}^{N} w_n \mathbf{x}_i^{\mathbf{T}} \mathbf{x}_n + \varepsilon_i$$

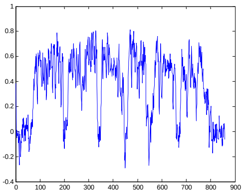
$$p(w_i) = N(0, \alpha_i^{-1})$$

$$p(\varepsilon_i) = N(0, \beta^{-1})$$

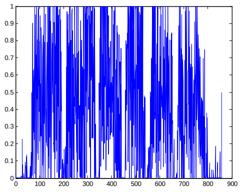
During the optimization, many elements of **α** will approach infinity. This means the corresponding values of **w** approach zero. The final result is a sparse **w**.

Constrained Quadratic Programming

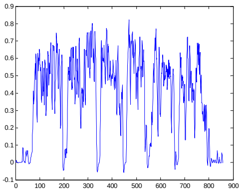
Quadratic programming was used to deconvolve the HRF from the predicted time courses, under the constraint that the solution falls in the range of zero to one. This constraint was imposed because of prior knowledge about the range of target scores. These solutions were then re-convolved by the HRF, and also by an additional Gaussian kernel (3 scans FWHM). Time limitations meant we did not attempt any more sophisticated Bayesian deconvolution strategies .



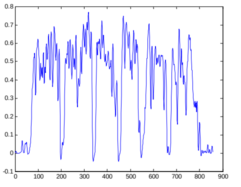
Original Prediction of movie 1, subject 1, language. Correlation 0.8577



Deconvolved data



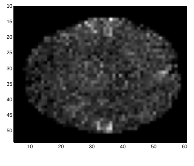
Reconvolved data correlation 0.8699



Smoothed data correlation 0.8854

Results

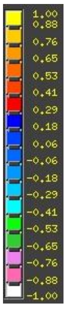
We ranked 5th in the competition. In general RVR performs slightly better than SVR, that is why we did not use SVR in the end. The results of combining 3 brains for predicting movie 3 of subject 1 and subject 2 improve the correlation. It is possible due to the averaging effect, as 3 subjects watch the same movies, and their rating may contain subjective errors. For the final submission of movie 3, subject 1 and subject 2 were predicted using training set with images of movie 1 and 2 for all subjects 1,2,3 with RVR, and their corresponding target ratings. But for subject 3, only images of subject 3 of movie 1 and 2 of are used.

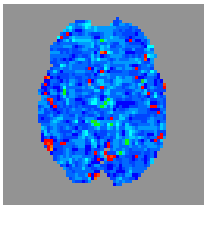


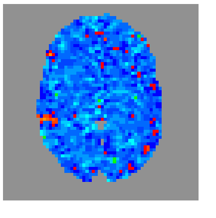
This image is the maximum intensity projection on the axial plane of the sum of weighted training images. It usually gives similar pattern as the SPM t-map projection.

$$\sum_{n=1}^{N} w_n \mathbf{x}_n$$

Sum of weighted training images







These 2 images are two particular slices in the axial orientation of the sum of weighted training images. These images are normalized to the range of -1 to 1.

Reference

[1] Friston KJ, Penny W, Phillips C, Kiebel S, Hinton G & Ashburner J (2002): Classical and Bayesian inference in euroimaging: theory.NeuroImage 16(2):465-483

[2] M. E. Tipping, Sparse Bayesian Learning and the Relevance Vector Machine, Journal of Machine Learning Research (2001), 1,211-244

[3] V. N. Vapnik, Statistical Learning Theory, Wiley, New York(1998)