Decoding fMRI with temporal integration: Learning the hemodynamical response function

M. Dyrholm[1], M.G. Philiastides[1]
R.I. Goldman[1,2], T.R. Brown[2], P. Sajda[1]
[1]Laboratory for Intelligent Imaging and Neural Computing
[2]Hatch MR Research Center
Columbia University, New York

Data analysis and Statistics:
Functional Neuroimaging
104.8 / BBB7

Introduction

There is a growing interest in employing multivariate methods for analyzing fMRI data, specifically as a way to exploit spatially distributed correlations linked to events/conditions of interest. Such approaches typically focus on learning spatial decompositions which optimize either a supervised or unsupervised objective function. However, fMRI is inherently a spatio-temporal signal and a principled approach should simultaneously find the spatial and temporal filters which optimize the objective of interest. [1] Bilinear logistic regression (BLR) has previously been applied for simultaneous learning of topographies and temporal envelopes in event-related EEG. [2] Here we present a version of BLR suitable for fMRI. The goal is to extract a spatial map of discriminating voxels and an associated hemodynamical integral for optimal inference about the experimental events (i.e. decoding).

Methods

Paradigm and data acquisition — Two subjects were scanned during sequential visual presentation of images from a database of human faces as well as random noise matched for mean luminance. Whole brain fMRI data were recorded on a 1.5T scanner (Philips Medical System) with gradient echo EPI, 24 slices of 64x64 voxels each, in-plane resolution of 3.125mm, slice thickness of 5.5mm, FOV=200mm, TE=40ms and TR=2s. For each subject, two different datasets were recorded:

1. Block design with enough scans to do robust simple averaging, suitable for localizing the FFA (c.f. Figure 1). Subjects passively viewed 12 alternating blocks of unmodulated, unmasked face and non-face images. Stimuli were presented for 750 ms with 250 ms inter-stimulus intervals, in blocks of 16 consecutive stimuli. A 12 s rest period was interleaved between blocks.
2. Interleaved design, suitable for trial-to-trial decoding (360 trials were recorded for each subject). A block of trials consisted of a total of 120 trials for both the face and non-face objects (60 each). Each subject performed a total of three blocks while we simultaneously recorded functional MRI data. Each image was presented for 100 msecs and it was followed by an inter-stimulus interval (ISI) that was randomized in the range 1-4 s (mean ISI=2.5s) in increaments of 250 ms. During the ISI subjects had to report whether they saw a face or a non-face object by pressing one of two buttons.

A square ROI defined to include the activated FFA was identified by averaging and subtracting the block design localizer scans for the two conditions (see Figure 1). This ROI is then detrended used to analyze the scans of the interleaved design. A temporal projection onto a Chebyshev basis of order 4 is subtracted in order to detrend and attenuate low frequency signal drift.

Event-related fMRI array nomenclature — To be able to learn an event-related response function, and to associate its (discrete) lags with temporal event-related latencies, we define the matrix $\mathbf{X}_n$ such that the element $(\mathbf{X}_n)_{x,\tau}$ is an estimate of the unobserved value of voxel-x at the defined latency δ_τ relative to the onset time of event n. The estimation is done by linear interpolation between the two acquisitions of voxel-x that happen to occur around event-n.

Single trial classification with Bilinear Logistic Regression (BLR) — We use the BLR algorithm of [2], and include a Laplace prior, to simultaneously learn a spatial and a temporal filter for maximizing discriminability between faces vs noise. BLR offers a physiologically motivated expansion of linear models to include multiple time lags

$$\mathrm{P}(\ \text{label}_n = \text{"face"} \mid \mathbf{X}_n\ ) \ = \ \psi\left[\ \sum_{r=1}^{R} \mathbf{u}_r^{\mathrm{T}} \mathbf{X}_n \mathbf{v}_r\ \right]\ ,\ \mathbf{u}_r, \mathbf{v}_r \sim \text{Laplace} \qquad (1)$$

The physiological motivation for factorizing space and time is the assumption of a common temporal envelope determined by the hemodynamical response function and associated spatial generators. The parameters $\mathbf{u}_r, \mathbf{v}_r$ are estimated on (360-fold) jackknifed surrogate data. The significance levels of the resulting parameter magnitudes are estimated by 1000-fold permutation test (where the null distribution is simulated by permuting the labels randomly). We assess the classifier accuracy by 10-fold cross-validated area under the ROC curve.

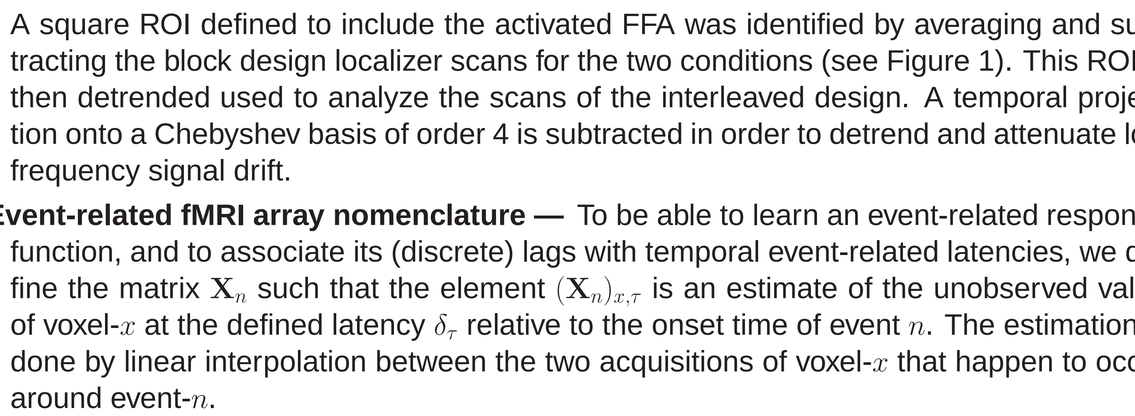

Figure 1: ROIs for the two subjects defined as described in the text. The ROIs are shown in color laid over the subject average scan in grayscale.

Results

The cross validated AUC for decoding the stimulus (discriminating face from noise trials) was **AUC = 0.90 (subject 1)** and **AUC = 0.79 (subject 2)** indicating robust decoding of the fMRI with this methodology. To represent the integrated space-time filter we form $\sum_r \mathbf{u}_r \mathbf{v}_r^{\mathrm{T}}$, and the temporal variation for this, at the spatial coordiate of the largest element of $\mathbf{u}_r$, is shown in figure 2. Note that the temporal filters qualitatively resemble the canonical hemodynamic response function, and the one for subject 1 includes an initial dip which is statistically significant at a 1% level (evaluated via permutation test).

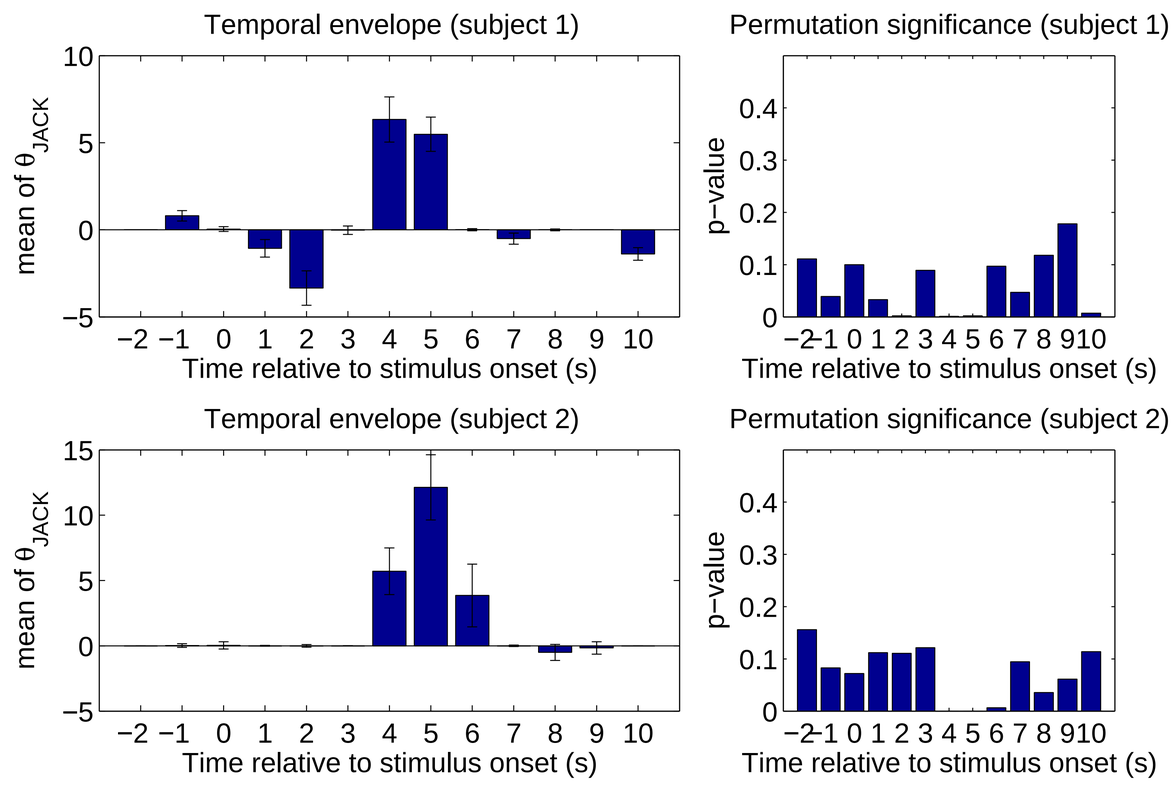

Figure 2: The jackknife parameter estimates, and magnitude significance (permutation) test result. Subject 1 shows a significant dip at a 2s latency after the event onset.

Discussion

We use a multi-variate machine learning approach to recover spatial and temporal filters which maximally discriminate stimulus conditions on a single-trial basis. Of particular interest is that for the face localizer task, the temporal filters that are learned by the model not only resemble the hemodynamic response, but that for some subjects this filter includes the controversial "initial dip" (at 2s). The fact the initial dip is recovered means that it contains discriminative information, implying a potential functionally significance, at least in some individuals. This initial decrease in the BOLD response is believed to arise from an increase in oxygen consumption and to be mostly microvascular [5, 6]. High field (7T) and high spatial resolution experiments in humans were able to find this initial dip with peak time ≈2 s after stimulation [7]. We conclude that our method can be used to uncover functional significant and distributed spatial and temporal activations in fMRI.

Acknowledgments

This work was supported by NIH Grant EB004730.

References

[1] J. Mourao-Miranda, K. J. Friston, and M. Brammer. Dynamic discrimination analysis: A spatial-temporal SVM. *NeuroImage*, 2007.

[2] M. Dyrholm, C. Christophoros, and L. C. Parra. Bilinear discriminant component analysis. *Journal of Machine Learning Research*, 8:1097–1111, May 2007.

[3] J.J. Marota, C. Ayata, M.A. Moskowitz, R.M. Weisskoff, B.R. Rosen, and J.B. Mandeville. Investigation of the early response to rat forepaw stimulation. *Magn Reson Med*, 41:247–252, 1999.

[4] A.C. Silva, S.P. Lee, C. Iadecola, and S.G. Kim. Early temporal characteristics of cerebral blood flow and deoxyhemoglobin changes during somatosensory stimulation. *J Cereb Blood Flow Metab*, 20:201–206, 2000.

[5] R.S. Menon, S. Ogawa, X. Hu, J.S. Strupp, P. Andersen, and K. Ugurbil. BOLD based functional MRI at 4 Tesla includes a capillary bed contribution: echo-planar imaging mirros previous optical imaging using intrinsic signals. *Magn Reson Med*, 33:453–459, 1995.

[6] X. Hu, T.H. Le, and K. Ugurbil. Evaluation of the early response in fMRI in individual subjects using short stimulus duration. *Magn Reson Med*, 37:877–884, 1997.

[7] E. Yacoub, A. Shmuel, J. Pfeuffer, P.F. Van De Moortele, G. Adriany, K. Ugurbil, and X. Hu. Investigation of the initial dip in fMRI at 7 Tesla. *NMR Biomed*, 14:408–412, 2000.