Choosing the optimal set of basis functions for fMRI data analysis using ROC methods with real data

R. R. Nandy[1], D. Cordes[2]

[1]Radiology, University of Washington, Seattle, WA, United States, [2]University of Washington, Seattle, WA, United States

Introduction

Any fMRI data analysis is most sensitive when we use an accurate model for the fMRI response. It is customary to model the HRF for a single impulse using some analytic functions. Popular choices include (a) a single gamma function to account for the hemodynamic delay [1], (b) difference of two gamma functions to account for the hemodynamic delay as well as an undershoot at the end of the impulse [2], (c) piecewise half-cosine functions to account for the hemodynamic delay as well as undershoots at the beginning and end of the impulse [3]. For a block design, the modeled HRF is convolved with a function representing the paradigm (typically a boxcar function). If the design is periodic, Fourier basis functions can be used as well. However, adding additional functions also can produce "nonsense functions" unrelated to the HRF. It is imperative that we draw an optimal balance between a better fit to the true response and overfitting the noise. However, the studies investigating the merits of each of these models have essentially been empirical in nature with no single objective criterion to choose an optimal model. We have recently introduced a novel Receiver Operating Characteristic (ROC) method using real fMRI data which can be used to establish a criterion to select the optimal set of basis functions [4]. We now describe how this new ROC method can be used to compare the response functions.

Methods

By definition, an ROC curve is a plot of True Positive Fractions (TPF) against False Positive Fractions (FPF) for all possible values of the test statistic used in the analysis. TPF cannot be estimated with sufficient accuracy from real fMRI data as we do not know which of the voxels among the detected voxels are truly active. However, it is possible to estimate FPF from real resting state fMRI data. If the subject from whom the fMRI data is acquired refrains from any active task, it is reasonable to assume the data to be pure noise with respect to the specific stimuli we are interested in. In such a case, all the voxels can be assumed to be inactive and FPF can be estimated for different values of the test statistic. However we already mentioned the difficulty in estimating TPF. Therefore instead of trying to estimate TPF, we estimate the fraction of all voxels detected to be active for different values of the test statistic. Clearly this includes both active and inactive voxels. Then it is possible to obtain a modified ROC curve using real data by using the fraction of detected voxels instead of TPF. At a given threshold for the value of the test statistic, active voxels are detected from the two collections and are respectively called Active Positives and Resting Positives as opposed to True and False Positives. We calculate the fraction of Active Positives (FAP) and the fraction of Resting Positives (FRP) and then plot the modified ROC curve. For the ROC method, we need two sets of fMRI data with identical scanner protocol, where the first set is pure resting-state data (the subject refrains from any active task with eyes closed) and the second set is the activation-state data. We analyze both datasets using the same set of basis functions and plot the fraction of active positives (FAP) against the fraction of resting positives (FRP) for different thresholds to plot a modified ROC curve. It is possible to reconstruct conventional ROC curves from the modified ROC curves [4].

This method can be applied to compare competing sets of basis functions. For each set of basis functions, we run our analysis on the resting-state data as well as the activation data and then plot the corresponding ROC curve. Once all the curves are plotted, we can compare the curves to choose the optimal set of basis functions. Since in fMRI, we are mostly concerned with thresholds at which FPFs are small, it may be more meaningful to plot restricted ROC curves for smaller values of FPF as opposed to the entire ROC curves running from (0,0) to (1,1). Also, it is important to observe that such an analysis is not possible using simulated fMRI data where the designated active voxels already have a specified HRF and by default, the same HRF will be the best choice for the post-processing and therefore cannot be a fair assessment of competing models for the HRF.

Results

As an example, we have used a dataset from a periodic visual encoding paradigm (30 sec encoding – 30 sec control, 5 periods) and a resting-state dataset, acquired with identical scanner parameters (FOV 24 cm x 24 cm, BW +/- 62.5 KHz, TR 1 sec, Flip 70 deg, slice thickness 4 mm/skip 0.5 mm, 64x64 resolution, coronal acquisition). The encoding task was to view and memorize novel pictures representing complex outdoor scenes (each image is displayed for 3 sec, 10 images total in 30 sec). The control task was to see the letters a or b, which were presented in random order. Both the datasets were analyzed using univariate regression for different choices of basis functions. In Figure 1, we have plotted the modified ROC curves for four different choices of basis functions; (i) a pure boxcar function with a 6 sec delay, (ii) a single gamma-convolved boxcar function (HRF1), (iii) difference of two gamma functions convolved with a boxcar function (HRF2), and (iv) Fourier basis
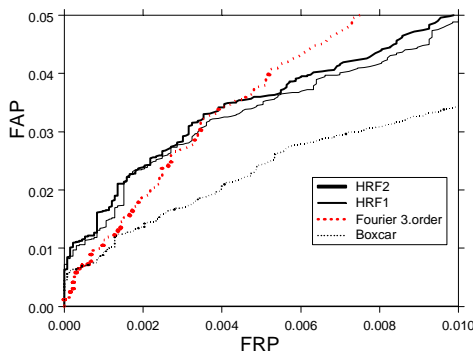
functions up to third order. The function HRF2 is the standard HRF model in SPM2. We have restricted the curves to values of FRP less than 0.01 for a better comparison, since the activation maps in fMRI always are produced using thresholds corresponding to very small values of FRP. Function HRF2 is the best choice for small values of FRP, closely followed by HRF1. This conforms to the fact that in a long block-design, the effect of the undershoot at the end of the impulse is not expected to be significant. The difference is likely to be more significant for an event related design. The Fourier basis functions perform worse than HRF1 and HRF2 for very small values of FRP, but for values larger than 0.004, the performance is better than HRF1 and HRF2. As expected, the simple boxcar function is the worst performer.

Figure 1. Modified ROC curves for four different basis functions: (i) a pure boxcar function with a 6 sec delay, (ii) a single gamma-convolved boxcar function (HRF1), (iii) difference of two gamma functions (HRF2, this is the standard HRF in SPM2) convolved with a boxcar function, and (iv) Fourier basis functions up to third order.

References

1. Boynton G, et al. J Neurosci, 16(13):4207-4221 (1996).
2. Friston K J, et al. Neuroimage, 7(1):30-40 (1998).
3. Woolrich M, et al. Neuroimage, 12: 1748-1761 (2004).
4. Nandy R and Cordes D. Magn Reson Med 49:1152-1162 (2003b).