Published in final edited form as:

Commun Stat Theory Methods. 2009 ; 38(16-17): 3099–3113. doi:10.1080/03610920902947576.

Estimation and Classification of BOLD Responses Over Multiple Trials

KUSH KAPUR^{1,3}, ANINDYA ROY^{1,2}, DULAL K. BHAUMIK^{1,4}, ROBERT D. GIBBONS^{1,4},
NICOLE A. LAZAR^{1,5}, JOHN A. SWEENEY³, SUBHASH ARYAL^{1,6}, and DAVE PATTERSON⁷

¹ Center for Health Statistics, University of Illinois at Chicago, Chicago, Illinois, USA

² Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, Maryland, USA

³ Center for Cognitive Medicine, University of Illinois at Chicago, Chicago, Illinois, USA

⁴ Division of Biostatistics, University of Illinois at Chicago, Chicago, Illinois, USA

⁵ Department of Statistics, University of Georgia Athens, Athens, Georgia, USA

⁶ Department of Biostatistics, University of North Texas Health Science Center, Fort Worth, Texas, USA

⁷ Discerning Systems Inc., Burnaby BC, Canada

Abstract

In this article, we model functional magnetic resonance imaging (fMRI) data for event-related experiment data using a fourth degree spline to fit voxel specific blood oxygenation level-dependent (BOLD) responses. The data are preprocessed for removing long term temporal components such as drifts using wavelet approximations. The spatial dependence is incorporated in the data by the application of 3D Gaussian spatial filter. The methodology assigns an activation score to each trial based on the voxel specific characteristics of the response curve. The proposed procedure has a capability of being fully automated and it produces activation images based on overall scores assigned to each voxel. The methodology is illustrated on real data from an event-related design experiment of visually guided saccades (VGS).

Keywords

Functional magnetic resonance imaging; Nearest neighbor; Spatial correlation; Spline; Temporal correlation; Wavelets

1. Introduction

As fMRI studies contain enormous amount of data, the analysis stream should entail statistically reliable yet computationally tractable methods for inference. Depending on the task presentation, fMRI studies can be broadly classified into block designs and event-related experiments. Block designs are based on repeated presentation of stimulus for several seconds followed by equal amount of rest period. Event-related experiments are composed of rapid presentation of stimulus with fixed or varying degrees of interstimulus

time. An experiment based on block design allows researchers to discern activation patterns, whereas event-related experiments provide them with a tool of exploring several BOLD characteristics such as strength of activation, percentage change, delay, rise rate, decay rate, etc.

In this article, our focus is to develop a simple and flexible method to analyze event-related experiment data incorporating both spatial and temporal trends. In order to do so, first we apply a 3D Gaussian spatial filter to reduce noise and borrow strength from the neighboring voxels in the data. Second, we use wavelets to model and remove some of the large scale features in the time series and thus detrend the data. Last component of our approach is the use of a fourth degree spline to fit the residual time course and to assign each voxel an activation score based on a set of classification rules discussed in Gibbons et al. (2004).

This procedure can be nearly automated and it can be easily extended to deal with variable length inter-stimulus intervals. Further, the specific advantage of analyzing multiple trials data is that we are able to incorporate voxel-specific variation in our analysis and measure the strength of activation in addition to the activation pattern.

Various authors have contributed to problems faced in event related designs; for example, Buckner (1998), Dale and Buckner (1997), Schacter et al. (1997), Glover (1999), Postle et al. (2000), Clark (2002), Gibbons et al. (2004). Buckner (1998), Dale and Buckner (1997), and Schacter et al. (1997) developed a method based on selective averaging of fMRI time course by trial type on each voxel and correlating it with posited stimulus to obtain activation maps using a test statistic derived from a t -distribution. Glover (1999) demonstrated the effectiveness of deconvolution in diminishing the hemodynamically imposed temporal blurring and quantifying the behavior of response in event related designs. Postle et al. (2000) modeled BOLD signal changes as convolution of postulated stimulus with impulse response function (IRF) for each subject. Using a general linear model framework they are then able to derive coefficients of interest from the convolved stimulus and observed data, and generate t -maps on voxel by voxel basis. Clark (2002) also used this framework to develop F -maps using a set of orthogonal polynomials upto the fifth order and used the Gaussian random field theory to deal with the multiple comparison problem. Gibbons et al. (2004) proposed a hierarchical polynomial regression model for simultaneously modeling and classifying the hemodynamic responses at multiple voxels for an event-related functional magnetic resonance imaging (fMRI) experiment. Specifically, they used a simple, but flexible, cubic model for the BOLD response within each voxel, where the data for each voxel in their method is reduced from $b(t + 1)$ observations taken equally spaced over time (b trials, each of length $t + 1$), to $t + 1$ observations by averaging over the b trials. Based on a mixed-effects regression model, they produced empirical Bayes estimates for the cubic approximations to the responses at each voxel, thus “borrowing strength” from neighbors. Properties of the empirical Bayes estimates, such as height and location of the maximum, location of the minimum, and the range of the estimated intensities, were then used to identify active voxels.

In our present work, we performed a detailed scrutiny of the empirical features of the entire time series and found that even though the activation patterns in trials are similar for an active voxel, the signal to noise ratio is very low and large-scale features between the trials, such as local drifts and breaks are confounded with the periodic nature of the trials. Under this scenario, it is no longer feasible to perform averaging over trials and apply simple polynomial models, since the estimates of the entire series will be discontinuous at the joining points of the trials and therefore, it can lead to a lower detection level in activation. Models for the response that impose continuity are appropriate in accordance with the Balloon model proposed by Buxton et al. (1998) to model dynamic changes in

deoxyhemoglobin content during sustained activity within a voxel for event-related designs. To this end, we use continuous spline models to approximate the underlying BOLD signal in each voxel. By modeling all b trials simultaneously we are able to produce graded activation images which provide a measure of strength of activation. We also propose a novel optimizing criterion which enables complete automation of our procedure with respect to the tuning parameters of the model.

We use wavelets in this article primarily to detrend the fMRI time series. Meyer and McCarthy (2001), Friman et al. (2004), and Bullmore et al. (2003) discussed wavelets techniques for detrending the fMRI time series data. Weaver et al. (1991), Angelidis (1994), and Zaroubi and Goelman (2000) used wavelets for denoising or whitening and image compression. Other uses have included multiresolution analysis of spatial maps of fMRI time series (Brammer, 1998; Desco et al., 2001), linear model estimation in the wavelet domain (Muller et al., 2003), and resampling in the wavelet domain, or “wavestrapping” (Bullmore et al., 2001, 2003). Whereas wavelets have a relatively long history in neuroimaging, splines have only recently begun to make an appearance in the fMRI literature. Carew et al. (2003) use smoothing splines to get an optimal smooth of the fMRI time series. They do this as a way of handling the autocorrelated residuals of linear models, that is, as an alternative to the whitening approaches described by Bullmore et al. (2003). Raz et al. (2003), on the other hand, use quadratic B-splines to estimate the hemodynamic response curve itself. These authors use splines as a modeling tool. Others (Macey et al., 2004; Tanabe et al., 2002) apply splines in the preprocessing stages of fMRI, to remove trends or global effects prior to statistical analysis. Friman et al. (2004) also discussed use of splines in the preprocessing of BOLD signals. Finally, spatial smoothing of the data is a well-accepted part of fMRI preprocessing.

In the next section, we outline the details of disparate elements of our analysis to create a model to capture both spatial and temporal trends for the fMRI data. We discuss the choice of the smoothing and thresholding parameters, introducing a criterion for evaluation based on two summary measures: the total number of active voxels and the number of singleton clusters (those containing a single voxel). Thus, the key technical concepts used in the paper are spatial smoothing, wavelets, spline fitting and classification rules. Results using a real data example are described in Sec. 3. We end with a discussion in Sec. 5.

2. Methodology

In this section, we describe our proposed methodology. The methodology will produce whole brain volume analysis of the fMRI data. The procedure has several steps and at the end it produces an activation score for each voxel. We will assume that there is a total number of V voxels in the brain and the index for each voxel will be denoted by v .

2.1. Step 0: Preprocessing with 3D Gaussian Spatial Filter and Wavelets

In the first step, a 3D Gaussian spatial filter is convolved with brain volume data at each time-point to capture the interdependence in contiguous voxels. Since the Gaussian filter is symmetric, the convolution operation for each voxel can be interpreted as computation of weighted average of the signal in each voxel with its surrounding neighbors. Therefore, this operation results in attenuated high frequency components and reduced noise levels along spatial dimension within the data. The large scale features of the BOLD response along the temporal direction can be eliminated by estimating and removing local trends. Most fMRI software programs perform this pre-processing step. Meyer and McCarthy (2001) showed how wavelets can be used to estimate local trends. Wavelets are flexible function bases that can approximate a variety of signal forms. For more details on wavelet estimation of local trends in fMRI signals we refer the reader to Meyer and McCarthy (2001). We tried a

variety of wavelet functions such as the Haar, Daubechis, Symmlets, Coiflets, and other wavelets available in the wavelet toolbox in MATLAB. The number of scales (resolutions) that one can use for a particular wavelet basis is limited by the finite length of the data. The highest level of multiresolution for the Symmlet8 analysis that was possible for the available length of the BOLD signal was 6. The level 6 resolution of the data was the smoothest scale approximation of the data and we used that for estimating the smooth underlying trend. Also, both Symmlets and Coiflets wavelets are nearly symmetric therefore wavelet coefficients well match to the original signal. Therefore, based on the approach described in Meyer and McCarthy (2001) and a direct visual inspection (see Fig. 1) we found that a coarse approximation by a smooth wavelet such as Symmlet8 can effectively estimate the local trends in the BOLD signal.

Figure 1 shows a typical BOLD signal along with the estimated trend for a data from a typical fMRI experiment. There is information in the average level of each voxel specific time series that may be relevant to the activation pattern. Because the estimated trend has the same mean level as the original series, information about the mean level of the series is lost after detrending. In order to retain the information about the mean level, the mean is added back to the detrended series. Thus, the observations used for modeling the signal from voxel v are

$$y_k(v) = x_k(v) - d_k(v) + \mu_x(v), \quad k=1, 2, \dots, b(t+1), \quad v=1, 2, \dots, V,$$

where $x_k(v)$ is the BOLD signal series obtained after Gaussian smoothing at voxel v $d_k(v)$ is the estimated drift at voxel v , and $\mu_x(v) = [b(t+1)]^{-1} \sum_{k=1}^{b(t+1)} x_k(v)$ is the mean value over all trials for the BOLD signal series at voxel v .

One motivation of the present article is to model any intra-trial correlation that maybe present in the BOLD signals. If the detrending followed by cubic fitting is able to adequately capture the general form of the entire b trial length time series, then any further refined analysis is unwarranted. To investigate any possible inter-trial dependence, we looked at the temporal correlation left in the residual time series obtained by subtracting the estimated cubic signal (repeated in each of the b trials) and a local trend from the original signal. We estimated the lagged autocorrelations for voxel v , $\hat{\rho}_v(h)$, for lags $h = 0, 1, 2, \dots, 120$. For each lag, the estimates were averaged over all voxels. Thus, the estimated autocorrelations are

$$\widehat{\rho}(h) = V^{-1} \sum_{v=1}^V \widehat{\rho}_v(h),$$

where

$$\widehat{\rho}_v(h) = [b(t+1)]^{-1} \sum_{k=h+1}^{b(t+1)} (y_k(v) - \mu_y(v))(y_{k-h}(v) - \mu_y(v)),$$

$$\text{and } \mu_y(v) = [b(t+1)]^{-1} \sum_{k=1}^{b(t+1)} y_k(v).$$

We found significant positive autocorrelation at lag one and at lag two in the residuals that can lead to loss in efficiency in estimation of the mean signal. Thus, care must be taken while summarizing the data at each voxel. A more refined analysis of the sequential information in the trials, if any, is warranted.

2.2. Step 1: Spline Fit

We propose to model the entire BOLD signal after the signal has been preprocessed and long term drift has been removed. In order to fit the BOLD responses for the entire time series we need to move beyond cubic polynomials. One can fit individual cubic models to each trial, but the estimate of the entire series will be discontinuous at the join points of the trials. Because the underlying signal is generally perceived as a continuous function one needs to impose the continuity assumption on the estimates as well. The most natural extension of the cubic model to the multiple trial case is the cubic splines model. Polynomial splines can adequately approximate many complicated functions by approximating segments of the function with lower order polynomials and joining the segments to form a continuous function. If smoother approximations are required, the polynomials in the different segments can be chosen to have a certain number of continuous derivatives at the joining points. For the case of BOLD response the segments are naturally selected as the basic trials associated with each run. Once continuity is imposed on the approximation, the behavior in one trial will depend on that in the adjacent trials. An artifact of this dependence is that in addition to the natural crests and troughs for active signals within trials, for certain trials a dip is observed at the beginning of the trial. A reason for this is that an unrestricted cubic model has four degrees of freedom (four free parameters) for estimating the BOLD signal in each trial and hence it can provide an adequate approximation. However, the number of free parameters for the spline model available for modeling within a trial is less than four and hence it does a poor job of capturing all the upturns and the downturns of the BOLD signal. Thus, to add flexibility, we fit piecewise fourth degree polynomials to the BOLD response with knots at the end of each trial. Therefore, the model for the BOLD signal at voxel v is:

$$y_k(v) = a_0(v) + \sum_{l=1}^b \sum_{j=1}^4 a_{j,l}(v) (k - b_l)_+^j + \varepsilon_k(v), \quad (1)$$

for $k = 1, 2, \dots, b(t+1)$, $v = 1, 2, \dots, V$, where the function $g(x) := (x - b)_+$ is defined as

$$g(x) = \begin{cases} 0 & \text{if } x \leq b \\ x - b & \text{otherwise,} \end{cases}$$

$b_l = (l-1)(t+1)$, $l = 1, 2, \dots, 3$ are the trial starting points which are the natural knots for the splines, and $\varepsilon_k(v)$ are assumed to be white noise with zero mean and constant variance σ_ε . For the sake of notational simplicity, we suppress the obvious dependence of the quantities on the voxels. Then for each voxel, the estimated signal is

$$\hat{y}_k = \hat{a}_0 + \sum_{l=1}^b \sum_{j=1}^4 \hat{a}_{j,l} (k - b_l)_+^j$$

where $\hat{a}_0$ and $\hat{a}_{1,l}, \hat{a}_{2,l}, \hat{a}_{3,l}, \hat{a}_{4,l}$, $l = 1, 2, \dots, b$ are the least squares estimates from the regression (1). Based on the spline fits $\hat{y}_k$ we produce activation scores for each voxel.

2.3. Step 2: Classification of Spline Fits

In an active voxel, the signal will peak to a significant height toward the middle of the trial and then dip to a low toward the end of the trial. Based on these features we can construct rules that will allow us to identify active voxels. For each voxel v we assign a score $s(v)$ which is equal to the total number of trials classified as active for that voxel. We can then construct a graded activation map for the whole brain on a scale of 0 to b based on the score $s(v)$. This allows us to grade our conviction about the activation of each voxel. For example, a voxel with a high value of $s(v)$ is more likely to be active. We can also construct a binary image of activation by thresholding the scores.

In order to assign scores we need to construct classification rules. We will use the rules given in Gibbons et al. (2004) with minor modifications. Suppose the maximum and the minimum of the estimated signal within each trial are CV1 and CV2 and the time point where the maximum and the minimum are reached are CP1 and CP2, respectively. Let t_1 and t_n denote the first and the final seconds within each trial. Let $\bar{x}$ denote the average BOLD response over all voxels, time points, and trials and let s_x be the corresponding standard deviation. The mean and the standard deviation of the detrended series, $\bar{y}$ and s_y are almost identical to those of the x series, and either one can be used in the classification rules without any significant loss of generality. Also, let $\hat{e}_i = \tilde{y}_i - \hat{y}_i$ be the residuals and let σ_e be the root mean square of the residuals from all the voxels, trials and time points. Then the response from a particular trial at a particular voxel will be classified as active if:

1. $CP1 \leq (t_n - t_1)/2$ and $CP1 > (t_n - t_1)/4$;
2. $CP2 < t_n$ and $CP2 > (t_n - t_1)(3/4)$;
3. $CV1 - CV2 \geq \sigma_e$;
4. $CV1 > \bar{x} - s_x$.

The only changes from those given in Gibbons et al. (2004) are that the time point $(t_n - t_1)/2$ is added to the range of possible time points for the maximum to occur and the requirement for a significant range within each trial has been relaxed by decreasing the threshold from $2\sigma_e$ to σ_e . The first change is due to the continuous nature of the spline estimates which may cause location of the minimum and the maximum within each trial to be more variable compared to when no such continuity is imposed. The reason for the second change is that the range of the fitted response in each trial is typically smaller than that when the fit is for the average over all trials. When the least squares fit are done over all trials simultaneously and the sum of squares of residuals over all trials is minimized at once, the peaks and dips in the estimated signal in each trial are less pronounced.

2.4. Step 3: Thresholding of Classification Scores

In order to convert graded images to binary activation map, one can simply threshold the graded activation map $s(v)$, $v = 1, 2, \dots, v$. Let τ be a universal threshold for the entire brain, then a voxel is called active if $s(v) \geq \tau$ otherwise it is deemed inactive. Here, τ is a percentage of a grade scale $[0, b]$. In practice, due to discreteness, the score $s(v)$ can only assume certain percentage values. However, the range of possible threshold values can be expanded by smoothing the raw scores. Thus, even though only a few meaningful values are possible for each application, we leave τ as a continuous measure, changing according to the scoring system. To implement the spline method the practitioner needs to choose values of the smoothing parameter (σ) and the threshold parameter (τ). In this article, we discuss few guidelines for choosing these parameters in an automated manner. The clarity of any activation image is generally summarized via the total activation percentage and the number of singleton clusters. The total activation percentage are high (within reasonable bounds) for rich activation patterns. The number of singletons is generally treated as misclassification.

We will investigate the quality of the activation patterns obtained for various choices of σ and τ by optimizing the following summary measures:

$$\begin{aligned} \text{TAC}(\sigma, \tau) &= \text{Percentage of active voxels, and} \\ \text{NSC}(\sigma, \tau) &= \text{Number of singleton active voxels.} \end{aligned}$$

We will maximize the richness of activation by maximizing TAC while having NSC has a penalty term. Figure 2 shows the percentage activation (TAC) as a function of the threshold τ for different values of the smoothing parameter σ for a typical fMRI experiment related to visual task performed at the UIC center for cognitive medicine studies. Each curve is a plot of TAC vs τ for a particular value of σ . The values of σ are taken from 0.6–1.2 per voxel unit with an interval of 0–2. Here, the distance between two adjacent voxels is considered as 1 unit. The range of τ depicted in the image is 5–30%. As expected, for a fixed σ , TAC is a decreasing function of τ . Note that the graphs are also quite smooth. The most encouraging fact that comes out of Fig. 2 is that the TAC graphs are close to each other and for moderately large values of τ the graphs get very close to each other. This indicates that the procedure is robust with respect to smoothing given that there is at least a moderate degree of smoothing.

Figure 3 shows the number of singletons (NSC) as a function of the threshold τ for different values of the smoothing parameter σ . Each curve is a plot of NSC vs τ for a particular value of σ . The values of σ are 0.6–1.2 in steps of 0.2. The range of τ depicted in the image is 5–30%. We see that the number of singletons is high with lesser smoothing, indicating a potentially inflated misclassification rate. But as σ is increased, the NSCs go down drastically and their values stabilize with respect to the threshold parameter.

The number of singleton clusters is in general low when there is some smoothing and the values shows remarkable stability for moderate to high values of σ . In practice, the value of σ can be taken to be anywhere between 0.8 and 1.2. We recommend $\sigma = 1$ as a default smoothing value for the brain volume acquisition parameters in our example. For a fixed value of σ , the activation images are functions of τ . TAC is a decreasing function of τ . NSC as a function of τ , although has multiple minima, definitely attains a consistency for σ greater than 0.8. If we let the threshold be too high, the number of active voxels will be zero and hence the number of singleton clusters will be zero. Also, if we let the threshold be too low, every voxel will be classified as active and hence no singleton cluster can form. Thus, NSC has two trivial minima at the two ends of the threshold scale. In between, there will be other local minima as new clusters can be formed and existing clusters can be broken up by changing the threshold continuously. Figure 4 shows the two quantities as a function of τ and σ .

For any given task, experts have a very good idea about the possible proportion of active voxels. For example, the percentage of active voxels in this task can be safely taken to be between 5 and 25%. Note that such bounds can be loosely defined without resulting in case by case choice. We would like to use this prior knowledge for choosing an optimal value of the threshold τ . We want to minimize the number of singleton clusters in the activation pattern with the constraint that the proportion of active voxels is between 5 and 25%. Suppose the prior opinion in any given slice is that between P_1 and $P_2\%$ of total voxels are active, then the objective criterion for finding the optimal value of τ is

$$\tau_{\text{optimal}} = \underset{P_1 \leq \text{TAC}(\tau) \leq P_2}{\text{argmin}} \text{NSC}(\tau). \quad (2)$$

Thus, the optimal value $\tau_{optimal}$ is the τ value which gives the minimum number of singleton clusters among all τ values for which the corresponding activation percentage is within P_1 and P_2 .

In practice, expert opinion about the bounds, P_1 and P_2 , of activation percentage may not be available for each new application and each new subject. However, the search for optimal value of τ can be conducted over a wide range of possible activation percentages. Thus, in absence of expert opinion we can designate a wide interval as possible values of activation percentages. Such a choice will be dictated by common-sense. Other measures that capture the quality of activation images and are more amenable to automation are a topic for future investigation.

3. Example

We illustrate our methodology and investigate the choice of threshold and smoothing parameters.

The data used in our example was an event-related design in which healthy control who gave informed consent to perform Visually Guided Saccades (VGS) task. The study was approved by the institutional review board. Each trial of the task began with 1 s of visually guided saccade followed by 14 s of fixation to be able to observe the hemodynamic response to the stimulus. Therefore, each trial resulted into a 15 point time series for each voxel where each point is associated with the corresponding second of the 15 s duration of the VGS task. Hence, $t = 14$ for our example. In total, there were $b = 23$ trials of this type which were retained by FIASCO preprocessing stream for motion correction. Each volume contained 15 slices in axial orientation, and each slice had 64×64 voxels. The visual targets moved every 750 ms. All imaging was performed on a 3-T whole body scanner (Signa VHi General Electric Medical Systems, Milwaukee WI) using the commercial quadrature radio frequency birdcage head coil. Acquisitions were performed with the commercial gradient echo, echo planar sequence (axial plane, TR = 1,000 ms, TE = 25 ms, FOV = 200×200 mm, matrix size = 64×64 , slice thickness = 3 mm). For our specific example, we have total of $V = 20,003$ in-brain voxels.

Figure 5(a) shows a typical BOLD response during a single trial. The general nature of the response in the active voxels can be described as follows: during a single trial of a task, after the stimulus has been presented, the response first rises and peaks, then falls to a trough as the oxygen level drops back to resting levels after the activation period. In some instances, the trough after recovery drops below baseline levels for a brief period as in the above example at 11 and 12 s. At the final time points in each trial the response recovers from the trough to a steady state level as oxygen level is restored. This behavior can be adequately characterized by a lower degree polynomial, specifically by a cubic function over the $t + 1$ time points.

Figure 5(b) shows the activation pattern in the third slice as a result of classification based on cubic hierarchical mixed-effects regression model of Gibbons et al. (2004). The primary areas of activation identified by the cubic classification method are the frontal and parietal eye fields. The frontal and parietal eye fields are the natural areas of activation for saccadic eye movements which was the task associated with this example; Luna et al. (1998). Thus, the activation pattern revealed by the cubic method is consistent with what one might expect with saccadic eye movements.

Figure 6 shows the spline fit and the cubic fit for two typical voxels. The spline fits the data much better than the cubic. This is not surprising since the spline uses many more parameters than the cubic. If the underlying model for the hemodynamic responses for the

active voxels were indeed a common pulse function (say cubic) which is repeated over trials, plus uncorrelated noise, then the quality of the cubic fit should be comparable to that of the spline. The superiority of the spline fit indicates that the averaging approach may fall well short of capturing features of the BOLD signal over a sequence of trials.

An advantage of the multiple trial methods is that one can have a finer resolution of the activation pattern. Figure 7 shows the brain activation during the fMRI experiment based on the raw scores obtained from the spline fit for 15 slices. The maximum value of the scores over all the voxels within the brain with smoothing parameter $\sigma = 1$ for isotropic Gaussian spatial filter, was 18 out of a maximum possible value of $b = 23$. Though there is a lot of noise in the brain (there is a high percentage of voxels with $s(v) = 1$), the principal areas of activation, the frontal and parietal eye fields, are brightly lit, indicating a high level of activation.

In order to create final binary activation images we threshold the classification scores. For the current example, we choose $P_1 = 5$ and $P_2 = 25$. Then the range of τ such that $5\% \leq \text{TAC}(\tau) \leq 25\%$ is $[4, 6]$. $\text{NSC}(\tau)$ starts attaining very low values when τ is in the range of $[3, 7]$ with σ greater than 0.8. For our analysis, we chose $\tau = 15\%$. Figure 8 shows the activation regions of slice 6 based on raw scores of the spline method. On the side of graded image, a comparison of activation is depicted between the spline method with $\tau = 15\%$ and the cubic classification. The activation pattern for the spline method is richer than that of the cubic in the sense that it reveals biologically relevant activation in more compact, well-connected clusters than those found by the cubic method.

Figure 9 shows the activation images of slice 6 for different values of the threshold τ when σ is held fixed at 1.0. From Fig. 9 it can be concluded that activation images associated with threshold values between 15 and 25% are conformable to general expert perception of anatomically correct activation in this particular slice for the given VGS task.

4. Discussion

In our article, we have explored several generalizations in the analysis of fMRI event-related experiments. The most important contribution of the present article is the introduction of the multiple trial paradigm. In the multiple trial method, the voxel-level data are taken to be the entire BOLD response series from consecutive trials rather than an average over the trials. The advantages of analyzing multiple trials data is that we are able to measure the strength of activation along with the activation pattern. We are also able to incorporate voxel-specific variation in our analysis. We also move away from slice by slice analysis and apply our methodology to analyze the brain as a whole. The end result of our procedure is a nearly automated method which gives anatomically more correct activation patterns and is better able to accommodate voxel-level variation. The strength of the proposed methodology is derived from its plurality. The present methodology can be extended to deal with variable length inter-stimulus intervals. The future work can extend on issues relating to the sensitivity and specificity of our method.

Acknowledgments

The authors thank two referees for their insightful comments that significantly improved the quality of this article. This work was supported in part by grant from the National Institute of Mental Health (1P50HD055751-01).

References

- Angelidis PA. MR image compression using a wavelet transform coding algorithm. *Magn Reson Imag.* 1994; 12:1111–1120.

- Brammer MJ. Multidimensional wavelet analysis of functional magnetic resonance images. *Hum Brain Mapp.* 1998; 6:378–382. [PubMed: 9788076]
- Buckner RL. Event-related fMRI and the hemodynamic response. *Hum Brain Mapp.* 1998; 6:378–382. [PubMed: 9788076]
- Bullmore E, Fadili J, Breakspear M, Salvador R, Suckling J, Brammer M. Wavelets and statistical analysis of functional magnetic resonance images of the human brain. *Statist Meth Med Res.* 2003; 12:375–399.
- Bullmore E, Long C, Suckling J, Fadili J, Calvert G, Zelaya F, Carpenter TA, Brammer M. Colored noise and computational inference in neurophysiological fMRI time series analysis: Resampling methods in time and wavelet domains. *Hum Brain Mapp.* 2001; 12:61–78. [PubMed: 11169871]
- Buxton RB, Wong EC, Frank LR. Dynamics of blood flow and oxygenation changes during brain activation: the Balloon model. *Magn Reson Med.* 1998; 39:855–864. [PubMed: 9621908]
- Carew JD, Wahba G, Xie X, Nordheim EV, Meyerand ME. Optimal spline smoothing of fMRI time series by generalized cross-validation. *NeuroImage.* 2003; 18:950–961. [PubMed: 12725770]
- Clark VP. Orthogonal polynomial regression for the detection of response variability in event-related fMRI. *NeuroImage.* 2002; 17:344–363. [PubMed: 12482088]
- Dale AM, Buckner RL. Selective averaging of rapidly presented individual trials using fMRI. *Hum Brain Mapp.* 1997; 5:329–340. [PubMed: 20408237]
- Desco M, Hernandez JA, Santos A, Brammer M. Multiresolution analysis in fMRI: sensitivity and specificity in the detection of brain activation. *Hum Brain Mapp.* 2001; 14:16–27. [PubMed: 11500987]
- Friman O, Borga M, Lundberg P, Knutsson H. Detection and detrending in fMRI data analysis. *NeuroImage.* 2004; 22:645–655. [PubMed: 15193593]
- Gibbons RD, Lazar NA, Bhaumik DK, Sclove SL, Chen HY, Thulborn KR, Sweeney JA, Hur K, Patterson D. Estimation and classification of fMRI hemodynamic response patterns. *NeuroImage.* 2004; 22:804–814. [PubMed: 15193609]
- Glover GH. Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage.* 1999; 9:416–429. [PubMed: 10191170]
- Luna B, Thulborn KR, Strojwas MH, McCurtain BJ, Berman RA, Genovese CR, Sweeney JA. Dorsal cortical regions subserving visually-guided saccades in humans: An fMRI study. *Cereb Cortex.* 1998; 8:40–47. [PubMed: 9510384]
- Macey PM, Macey KE, Kumar R, Harper RM. A method for removal of global effects from fMRI time series. *NeuroImage.* 2004; 22:360–366. [PubMed: 15110027]
- Meyer, FG.; McCarthy, G. Estimation of baseline drifts in fMRI. In: Insana, MF.; Leahy, RM., editors. *Information Processing in Medical Imaging. Lecture Notes in Computer Science.* London, England: Springer Verlag; 2001. p. 232-238.
- Muller K, Lohmann G, Zysset S, von Cramon DY. Wavelet statistics of functional MRI data and the general linear model. *J Magn Reson Imag.* 2003; 17:20–30.
- Postle BR, Zarahn E, D'Esposito M. Using event-related fMRI to assess a delay-period activity during performance of spatial and nonspatial working memory tasks. *Brain Res Protoc.* 2000; 5:57–66.
- Raz J, Zheng H, Ombao H, Turetsky B. Statistical tests for fMRI based on experimental randomization. *NeuroImage.* 2003; 19:226–232. [PubMed: 12814573]
- Schacter DL, Buckner RL, Koutstaal W, Dale AM, Rosen BR. Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. *NeuroImage.* 1997; 6:259–269. [PubMed: 9417969]
- Tanabe J, Miller D, Tregellas J, Freedman R, Meyer FG. Comparison of detrending methods for optimal fMRI preprocessing. *NeuroImage.* 2002; 15:902–907. [PubMed: 11906230]
- Weaver JB, Xu YS, Healy DM, Cromwell LD. Filtering noise from images with wavelet transforms. *Magn Reson Med.* 1991; 21:288–295. [PubMed: 1745127]
- Zaroubi S, Goelman G. Complex denoising of MR data via wavelet analysis: Application for functional MRI. *Magn Reson Imag.* 2000; 18:59–68.

Figure 1.
A typical BOLD response over 39 trials and the corresponding trend estimate.

Figure 2.
TAC as a function of τ for different values of σ .

Figure 3.
NSC as a function of τ for different values of σ .

Figure 4. Total activation percentage and number of singleton clusters as a function of the threshold τ and σ .

Figure 5.
A typical BOLD response and activation pattern after cubic classification.

Figure 6. Spline and cubic fit to detrended BOLD signals from two voxels. The estimated trends have been added back to the fits to plot them in the scale of the original signal.

Figure 7.
Different grades of activation based on classification scores.

Figure 8.
Regions of activation.

Figure 9.

Activation after thresholding $\hat{s}(v) > \tau$. The images are arranged from left to right and then top to bottom corresponding to values of τ from 5–30%.