Chapter 2
First-Level fMRI Data Analysis for Activation Detection

Abstract As many other scientific experiments, functional magnetic resonance imaging (fMRI) study has multilevel hierarchical nature. For example, for a single voxel in the brain, fMRI response is from *i* voxel position, in *j* run, of k subject. Therefore, a mixed effect model should be employed for the analysis. In framework of general linear mixed model (GLMM), both first and higher levels of the model parameters can be estimated using iterative schemes for the fMRI study. However, this is not practical for fMRI data analysis because of the computational burden of fitting GLMM at every voxel that is very expensive. As a result, most methods adopt a two-stage model method, and at the first stage, variance and effect from the first-level analysis are estimated for the second-/higher-level fMRI data analysis. In this chapter, we will focus on these variance and regression parameter estimation. Firstly, we present the method for fMRI experimental design, and then we introduce statistical methods for fMRI activation detection. A general linear model (GLM) for the activation detection is introduced. Because the residual terms of the linear model is autocorrelated, an autoregression (AR) model is adopted for correcting the model estimation error. After the fMRI signal is fitted with the GLM-AR model, we detail statistical inference method for activation detection. Finally, we present false discovery rate (FDR) and family-wise error (FWE) algorithms for the threshold correction.

Keywords BOLD-fMRI • fMRI experimental design • General linear model (GLM) • Autoregression (AR) • Threshold correction • Activation detection • Matrix inversion

Contents

X. Li, *Functional Magnetic Resonance Imaging Processing*,
DOI 10.1007/978-94-007-7302-8_2,

2.1 fMRI Experimental Design

The first step for processing any functional data is to understand the experimental design, and fMRI data analysis is no exception. In fact, a good fMRI experimental design is essential for the study since it not only provides the interpretable results of the scientific questions we want to address but also offers easier way to analyze the data. Because blood-oxygen-level-dependent (BOLD) response relies on the sluggish cerebral blood flow, this makes the functional magnetic resonance imaging (fMRI) experimental design different and, certainly, much harder than other types of experimental design. Therefore, many theories such as statistics, mathematics, and system theories have been applied to design fMRI experiments and analyze fMRI data. For example, if we regard the human brain as a system [1, 2], then the purpose of experimental design is to devise a system input. Since we know the predefined system input and system output, i.e., BOLD-fMRI response, we can infer what is inside the brain in response to this particular stimulus (system input). The purpose of fMRI activation detection is, therefore, to find the significant relationship between system input (experimental design or external stimuli) and fMRI response (system output). Although there are many types of fMRI experimental design available, three types of designs are commonly in use for fMRI studies currently, i.e., block design, (random) event-related (ER) design, and phase-encoded design. We will introduce these designs briefly in this section.

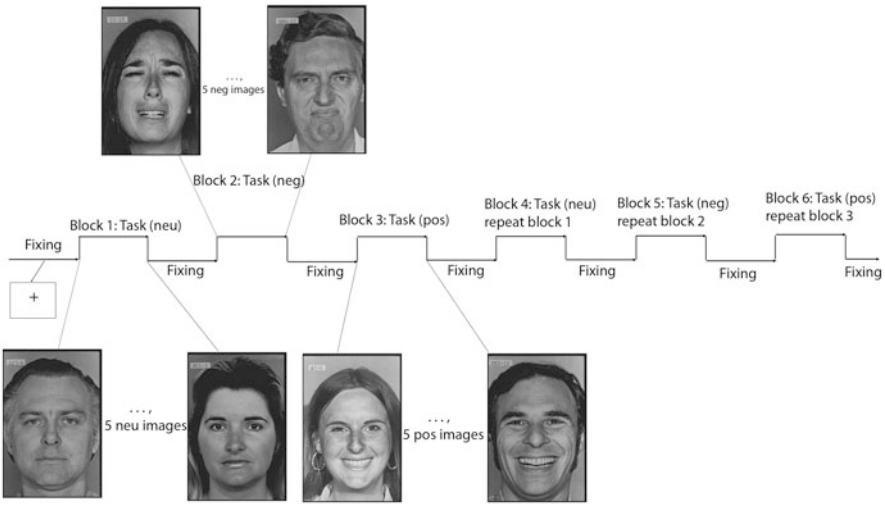

Fig. 2.1 An example of block design: emotional face processing experiment. One example of standard block design paradigm where on and off length of each block is equal: pos, neg, and neu denote positive, negative, and neutral face, respectively. Fixing is eye fixation to the center of the screen

2.1.1 Block Design

Perhaps, block design is the simplest and most frequently used experimental design for fMRI study. In a typical block design, there is a control condition block with one or more task blocks, and the control block and task block appear alternatively. The duration of each block is about 10–20 s. The length of these blocks can be equal or unequal. If the length of each condition block is equal, then it has additional advantage, i.e., Fourier analysis method can be applied straightforward because the stimuli change periodically.

Figure 2.1 shows one example of block design for emotional face processing. In this design, there is one white cross fixing control condition and three face/task conditions, and each block appears alternatively. The task conditions include neutral faces, sad faces, and happy faces. There are five faces with the same type of presentation within each block, and the duration of each face is 2 s.

In this experimental design, we can study the brain response to the face stimuli by comparing all the face condition response with the white cross fixing condition. In addition, we are able to address the emotional effect by comparing happy/sad face response with the neutral face condition. Finally, the question of the brain region difference in response to happy and sad faces can also be answered using statistical methods.

Ideally, it is always better to collect as many brain fMRI responses as possible. However, because subjects cannot maintain their positions in the scanner for a long time, we often take each functional run last 6–20 min. If more fMRI time slots are available, we can employ repeated run measure methods to collect more data from the same subject.

It should be mentioned that during the fMRI experiments, subjects are asked to perform certain task which may be irrelevant to stimuli but to help subject to keep awake in the scanner.

2.1.2 Random ER Design

Another type of fMRI experimental design commonly in use is ER design. It has been used for many studies including, but not limited to multisensory, visual/auditory psychophysics, perception learn, and neuronal adaptation studies. As the name indicated, in the random design, each event or condition/block will appear randomly so that the subject cannot guess what the next stimulus is during the experiment. It is worthwhile to mention that the rapid ER design with randomly select interstimulus interval belongs to ER design, and it has been becoming popular for fMRI studies. When using this type of experiment design paradigm, because hemodynamic delay of the cerebral blood flow, it should be kept in mind to leave enough time interval between two events [3].

One example of random ER design paradigm which was used for emotional face processing study is shown in Fig. 2.2. In this ER design, we include four conditions, e.g., three task conditions (neutral, sad, happy face stimuli) and one control stimulus (white cross fixation) condition. The presentation of each task condition is randomly selected, and there is a white cross eye fixation control condition after each neutral, sad, and happy faces stimuli condition. The duration of white cross is varied according to the duration of the face stimulus. For instance, if a single face which takes 2TR (TR is the scanner repetition time, in this example experiment, we set TR = 2 s) to present, a white cross stimulus with 4TR duration is displayed subsequently. We represent this task as event block 1 in Table 2.1, and it takes 6TR to finish the task in total.

The second event block in Table 2.1 is a double face stimuli (same type of the stimulus presented two times, i.e., both happy/sad/neutral faces, but different person face) presentation, which needs 4TR (2TR + 2TR) for faces presentation and 6TR for white cross displays.

The third type of event block is a triple stimuli presentation, which includes 6TR (2TR + 2TR + 2TR) for the same type of face presentation and 7TR for white cross display. The role of the white cross is to separate face condition with other types of face stimuli tasks.

For each condition and each case, we have three repetitions. It is easy to see that the design is balanced, i.e., the total number of each condition is the same.

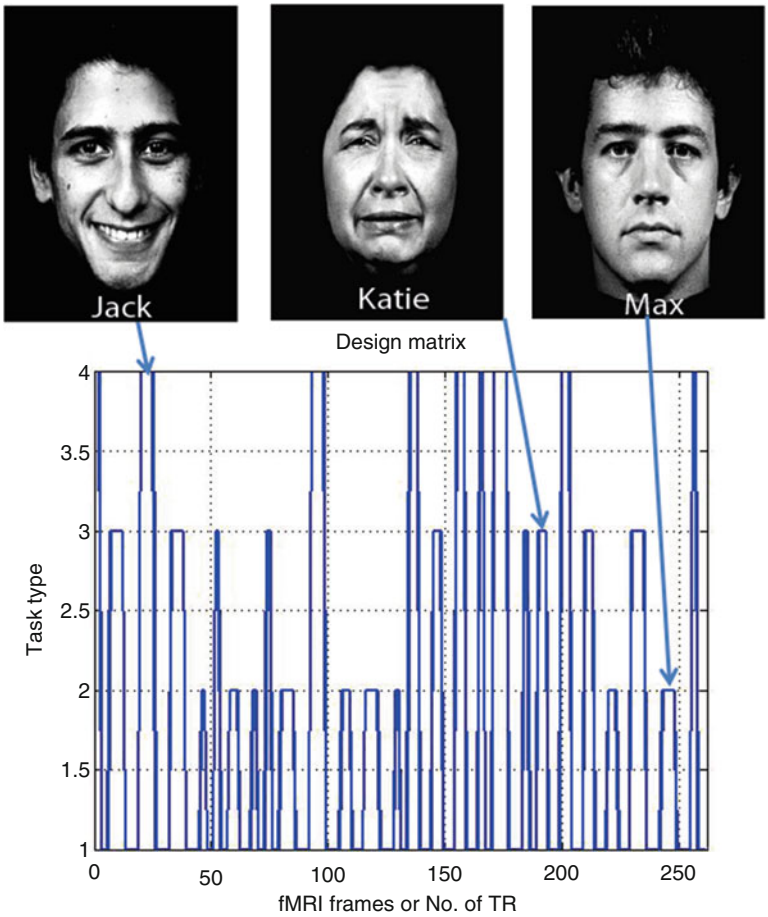

Fig. 2.2 An example for random ER design paradigm. *X*-axis is time (unit TR). In the *y*-axis, *1*, *2*, *3*, *4* denote white cross fixation, neutral, sad, and happy face stimulus presentation, respectively

Table 2.1 Task assignment for random ER emotional facial processing study

Face type	Neutral	Sad	Happy	No. of TR(2 s)
Event block 1	1,2,3	4,5,6	7,8,9	$9 \times 6 = 54$
Event block 2	10,11,12	13,14,15	16,17,18	$9 \times 10 = 90$
Event block 3	19,20,21	22,23,24	25,26,27	$9 \times 13 = 117$TR
Total time				261TR = 8 min 42 s

We arrange each event from 1 to 27 as shown in Table 2.1, and then we randomly select each number; as a result, each event block will appear randomly. Therefore, the interval between each even block is randomly determined accordingly. The total time for this experiment is 8 min 42 s, as given in Table 2.1.

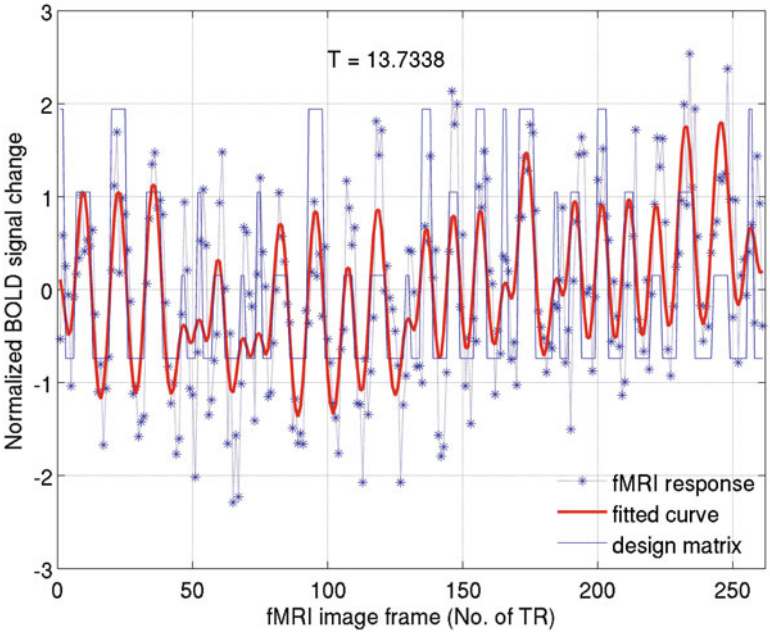

Fig. 2.3 An example of fMRI response for random event-related (*ER*) block design for emotion study. *T* is T statistical test

We may want to study the correlation between brain activation and behavioral task; we need to record the task performance during the MRI scan for the analysis. In addition, we may want to do additional behavioral tests after MRI scan, for example, in this experiment, during the fMRI experiment in the scanner, and we ask the subject to remember the face and corresponding name for a memory test after MRI scan.

To help understand the experimental design, we give one typical fMRI response (Fig. 2.3) from the experimental design paradigm as shown in Fig. 2.2. In Fig. 2.3, the blue thin solid curve denotes the normalized experimental design in Fig. 2.2, the dotted curve represents the fMRI response, and the red thin curve displays the model fitting for the response when comparing faces with white cross fixation condition (data is from Appendix D). It is obvious that the fMRI response follows the experimental design, suggesting that there is activation at this particular brain voxel position.

2.1.3 Phase-Encoded Design

Another regularly used fMRI experimental deign is phase-encoded design. In contrast to block and ER designs, although phase-encoded design is less frequently

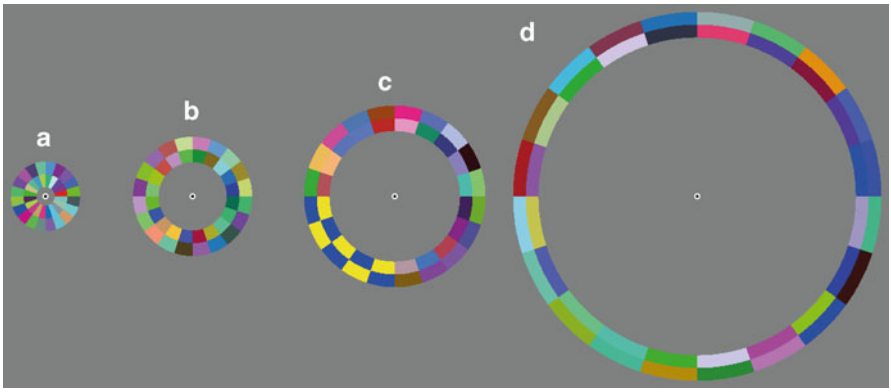

Fig. 2.4 Phase-encoded expanding ring stimulus for visual cortex retinotopic mapping study. (**a–d**) Ring expanding from small (**a**) to big (**d**)

in using, it has been applied successfully in visual retinotopic mapping studies [4–6] and proved to be another effective way to study visual cortex.

The idea of this type of design is to present a travelling wave (shift block) type stimulus in which stimulus changes continuously, without obvious interruption. For example, in visual retinotopic expanding ring stimulus, we change the position of the checkerboard pattern slowly over time, a flicker checkerboard color ring stimulus growing from small size (a) to big (d) as shown in Fig. 2.4. Then another small ring appears in the central of the stimulus to replace the largest ring stimulus in Fig. 2.4d, and the whole process repeated itself. This is the eccentricity stimulus for retinotopic mapping study. We also can create a contracting ring, in which the size of the ring changes from biggest (d) to smallest (a). Combining with phase-encoded wedge (polar angle) stimuli (Fig. 2.5), we can calculate the visual sign map and define the boundaries of the early visual cortex accordingly [4–6].

Figure 2.5 shows the polar angle stimulus for the retinotopic mapping study. Typically, it consists of a rotated wedge, whether clockwise or anti-clockwise rotation. When doing the experimental in the scanner, the subjects are asked to keep eye fixation at the center of the stimulus and perform certain task, i.e., detecting the regular black/white (Fig. 2.5a), blue/yellow (Fig. 2.4c), and red/green (Fig. 2.5b) patterns in the stimulus.

Based on the idea of phase-encoded design, we extended the ring stimuli to spatial frequency [7] and contrast sensitivity perception [8] stimuli for visual cortex studies. For instance, in the phase-encoded spatial frequency design, the spatial frequency changed periodically either from medium to low or from low to medium (Fig. 2.6) in which the spatial frequency of a sinusoidal checkerboard stimulus was gradually varied from 0.5 to 6 c/d over a 1 min period. The temporal frequency of the checkerboard stimulus was 8 Hz. This involved a smooth and gradual change in the spatial frequency of the sinusoidal checkerboard evenly throughout the field. A central fixation point was provided. The stimulus and design matrix are depicted in Fig. 2.6c. The attention of the subjects was controlled using a target detection

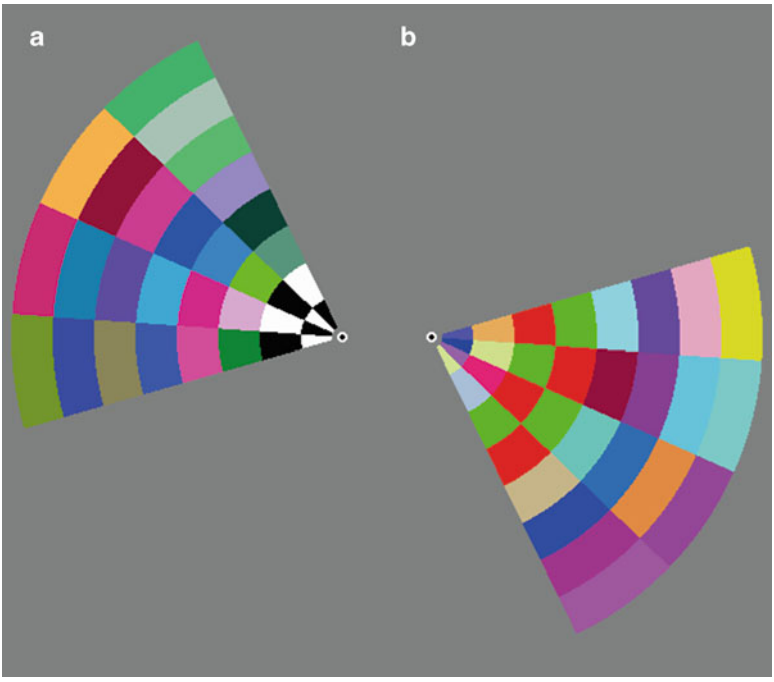

Fig. 2.5 Phase-encoded polar angle (wedge) stimulus for visual cortex retinotopic mapping study. (**a-b**) Different angle of rotated wedge stimuli

task as described above for the block design. The advantage of this design is that the stimulus changes smoothly, rather than dramatically changes as in block design, and the fMRI response will change smoothly accordingly which is easier to deal with using Fourier analysis method. Because from Fourier analysis, we know that the periodical triangle stimuli (Fig. 2.6c) can be approximated by a set of sinusoidal functions. The fundamental frequency sinusoidal function of the triangle stimulus can be employed to model the fMRI response (Fig. 2.7). In Fig. 2.7, the dotted curve indicates the fMRI response (data is from Appendix E), and the solid curve denotes the fundamental frequency of the triangle wave (Fig. 2.6c) as a model for hemodynamic response function (HRF).

2.2 fMRI Data Preprocessing

After fMRI data have been collected, we need to preprocess the data to correct subject head motion, register individual subject brain image to a standard template for comparison, and normalize the fMRI time series to remove the position shift in the response.

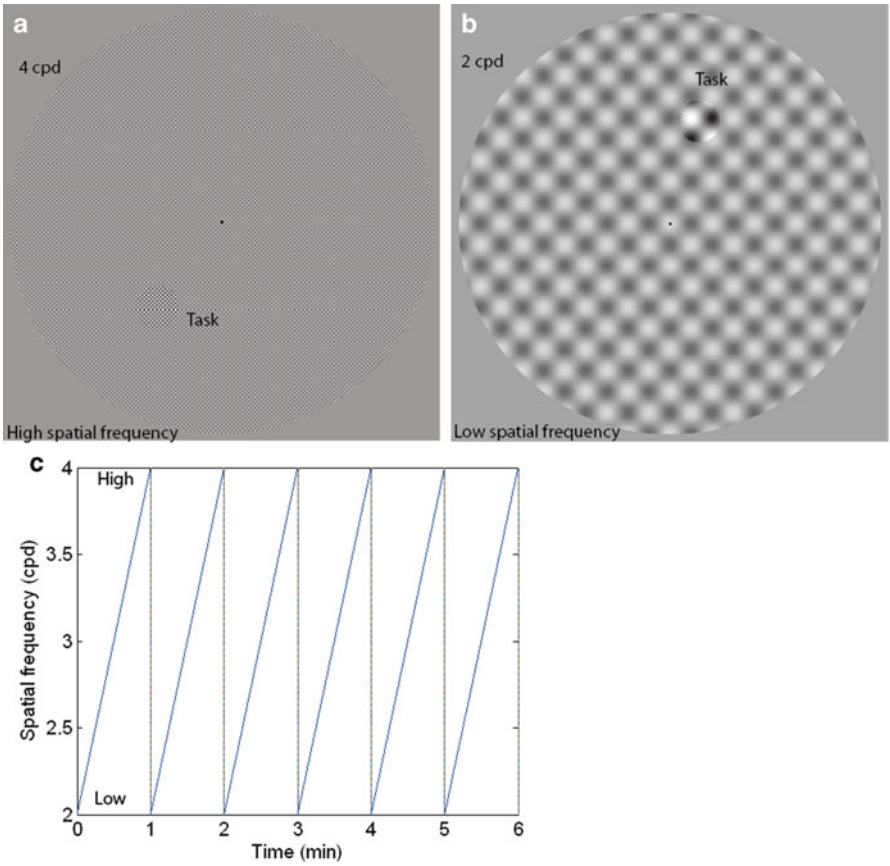

Fig. 2.6 Phase-encoded design for spatial frequency perception. (**a**) An example of high spatial frequency stimulus. (**b**) An example of low spatial frequency stimulus. (**c**) The stimuli presentation paradigm. *cpd* denotes cycle per degree

2.2.1 fMRI Data Motion Correction

In a typical fMRI experiment, we collect not only functional data but also structural MRI data for brain activation localization, visualization, and comparison with other subjects. We may also want to compare different subject from different image modalities; as a result, we need to register different subjects to a standard template for comparison. To achieve these goals, we need to register functional data to a high spatial resolution structural image, which needs to be matched to the standard template for group comparison. Image registration methods [9, 10] for these purposes have been developed, and there is a lot of software packages freely available.

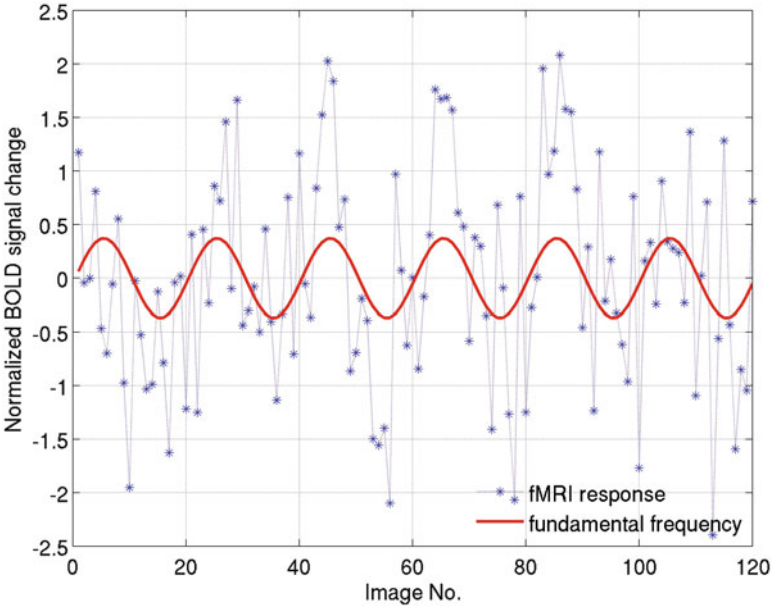

Fig. 2.7 An example of fMRI response to the phase-encoded spatial frequency experimental design

For functional images, we perform image registration to correct head movement during data collection within each run and across different runs. In the literature of image registration, many methods have been developed. For example, correlation method, linear and nonlinear least squares methods, and mutual information methods have been developed. In our study for fMRI data preprocessing, dynamic motion correction for functional image time series for each run and for different runs was realigned at the same time by using the fmr_preprocess within MINCtools (http://www.bic.mni.mcgill.ca/ServicesSoftware/HomePage) with default parameters for three-dimensional Gaussian low-pass filtering [11].

2.2.2 fMRI Time Series Normalization

After motion correction, the next step is to normalize the fMRI time series to exclude the baseline shift in the analysis. There are two ways to normalize fMRI time series, i.e., temporal and spatial normalization methods. For the spatial normalization, we need to segment the brain region from the image background and using the voxels within the brain region for normalization. For the temporal normalization, we normalize the fMRI response longitudinally, i.e., if we employ x_i to denote the fMRI signal at time point (fMRI image frame) i, then we use

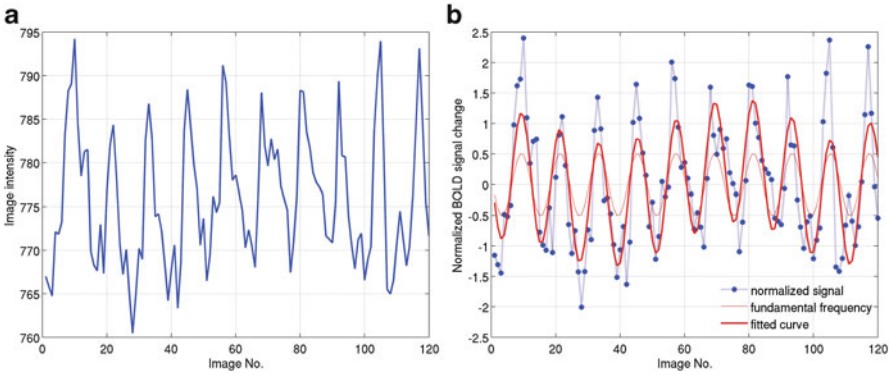

Fig. 2.8 Temporal normalization of fMRI response for a standard block design. (**a**) Raw fMRI image intensity time series. (**b**) Normalized fMRI signal

$$Y = \frac{x_i - \overline{x}}{\text{std}(x)} \tag{2.1}$$

to normalize the signal, where $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the mean value of the time series and $\text{std} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}$ is the standard deviation of the time series.

Figure 2.8a shows one example of fMRI image intensity time series from previous study [12] which adopted a typical standard block design (see Appendix F for data collection). Figure 2.8b shows the corresponding temporal normalization results using Eq. (2.1). The thin solid curve is the fundamental frequency model for HRF, and the thick solid curve denotes the fitted results. Comparing Fig. 2.8a with dotted curve in Figure 2.8b, it is easy to see that although the magnitude of these two curves is different (the baseline in Fig. 2.8b has been removed), the shape is similar, suggesting fMRI normalizing step maintains the temporal information of the signal changes. In Figs. 2.3 and 2.7, the solid curves represent the temporal normalized fMRI responses.

2.3 Activation Detection: Model-Free and Model-Based Methods

After fMRI time series have been corrected for head motion and normalized, it is ready for activation inference using statistical methods. Methods for fMRI activation detection can be classified into two broad categories: one is the model-free method and the other is model-based method. For the model-free method, we do not need

to build any fMRI response models, i.e., HRF beforehand for activation detection. While for the model-based method, we need to build a model for the analysis. In this chapter, we will introduce both model-free and model-based methods for activation detection at first-level analysis.

2.3.1 Model-Free Method: Two Sample t-test for Activation Detection

The simplest model-free method for activation detection is to compare the task fMRI response with the control condition response. The idea of this method is to compare the mean value of the neuron population response from task condition and control condition. This can be done based on the assumption that the task stimulus response and control condition response are two independent variables, randomly sampled from an approximately normal distribution. Therefore, we can apply a two-sample t-test as [13]

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{2.2}$$

for activation detection, where variance s can be calculated as

$$s = \frac{\sum_{j=1}^{n_1}\left(x_j - \overline{x}_1\right)^2 + \sum_{i=1}^{n_2}\left(x_j - \overline{x}_2\right)^2}{n_1 + n_2 - 2} \tag{2.3}$$

where $\overline{x}_1$ and $\overline{x}_2$ are the mean values of task stimulus response and control condition response. The degree of freedom (df) for the hypothesis test is $k = n_1 + n_2 - 2$. n_1 and n_2 are the total number of time point from the task condition and control condition, respectively. The next step is to compare the calculated T value with k df with the critical t value from the t distribution table at the chosen confidence level and decide whether to accept or reject the null hypothesis. Reject the null hypothesis when the calculated T value is larger than the critical T value, e.g., declare activation at this voxel.

The major advantage of this method is simple, as the t-test method does not require any prior defined HRF model for activation detection. It only needs the information of the stimuli duration and starting time. However, this method is based on the assumption that the individual fMRI signal is i.i.d., and this is not realistic because of the nature of the hemodynamic response of the blood flow. Moreover, this method does not take fMRI response slow drift into account, which can lead to estimation error.

2.3.2 Correlation Analysis Method

Correlation analysis is another model-free method for activation detection [14]. To apply correlation analysis method for activation detection, one needs to select a seed point time series which is supposed to be activated for this particular brain region in response to the stimulus. Then we correlate this fMRI time series from the predefined seed region with the rest of the brain. We obtain the correlation coefficient (CC) and use CC to determine whether it is significantly correlated or not. We can also covert CC to Z score and t value for significant inference. Pearson correlation coefficient is often adopted for the calculation as

$$r = \frac{\sum_{i=1}^{n} (x_{1i} - \overline{x}_1)(x_{2i} - \overline{x}_2)}{\sqrt{\sum_{i=1}^{n} (x_{1i} - \overline{x}_1)^2 \sum_{i=1}^{n} (x_{2i} - \overline{x}_2)^2}} \tag{2.4}$$

where x_{1i} is the seed region time series and $\overline{x}_1$ is the mean value of the fMRI series. It can also take the form of averaged time series from a region; x_{2i} is the fMRI time series from the rest of the brain, $\overline{x}_2$ is the mean value of this series, and n is total number of time series.

The advantage of this method is that it is simple to implement; however, this method is also easy to be disturbed by the slow drift, especially, in the seed region. Moreover, if the seed region is not selected properly, this can lead to large estimation error for activation detection. Although linear and quantic shifts have been incorporated to eliminate the slow drift effect, brain activation detection using this method is highly dependent on the predefined seed region. Human cerebral anatomy knowledge is also needed to decide the seed region/voxel.

2.4 Models for Hemodynamic Response Function and Drift

Because of the limitations of model-free methods for fMRI activation detection, model-based method such as general linear model (GLM) has been developed and used for activation detection. The first step to apply model-based method is to build an HRF model for the fMRI response. Currently, there are mainly three HRF models for fMRI activation detections, i.e., fundamental frequency of Fourier component model, two-gamma function model, and boxcar function convolve with Gaussian function model.

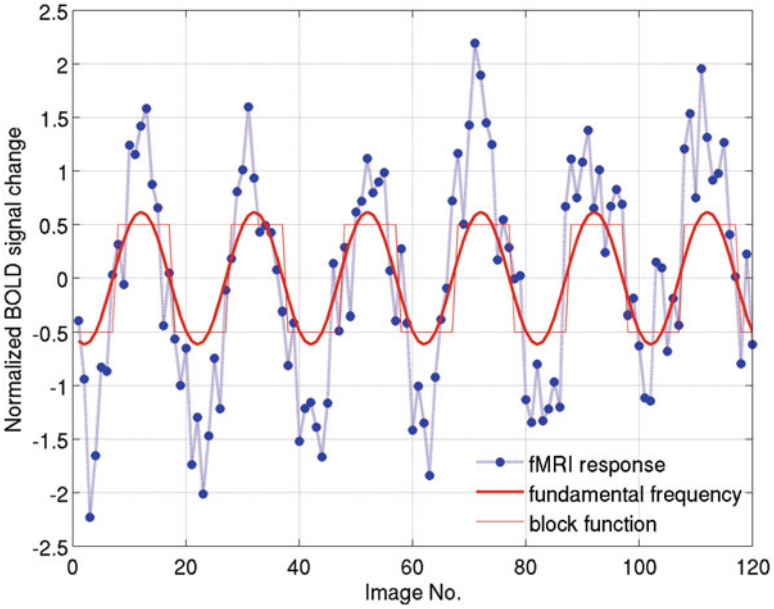

Fig. 2.9 One typical fMRI response and its fundamental frequency model for HRF from phase-encoded retinotopic mapping study. The block function (*thin solid curve*) and the fundamental frequency of the response (*thick solid curve*); the estimated onset for the first block is 8 from the FFT analysis

2.4.1 HRF Models for Activation Detection

If the stimuli in the experimental design are changing periodically and the duration of each block condition is equal (Figs. 2.7, 2.8b, and 2.9), then fast Fourier transformation (FFT) method can be applied to model the HRF from fMRI response [15]. This is because the periodical square block wave can be approximated by the fundamental frequency of the sinusoidal wave (e.g., Fig. 2.8), and it is a powerful way to study experimental design with periodical change stimuli. The sine wave function has the form

$$f_{t,1} = a \cos(\omega t + \theta) \tag{2.5}$$

where θ is the delay/onset or phase of the response that is estimated using the FFT method; ω is the angular frequency; $\omega = 2\pi f$, where f is the frequency of the stimulus; and a is the magnitude. The algorithm to estimate parameters θ and a in Eq. (2.5) is as follows:

1. Calculate fast Fourier transform (*fft*) of the normalized fMRI signal (e.g., dotted curve in Figs. 2.7 and 2.8b, or Fig. 2.9).
2. Get the angle and magnitude of the signal from FFT.

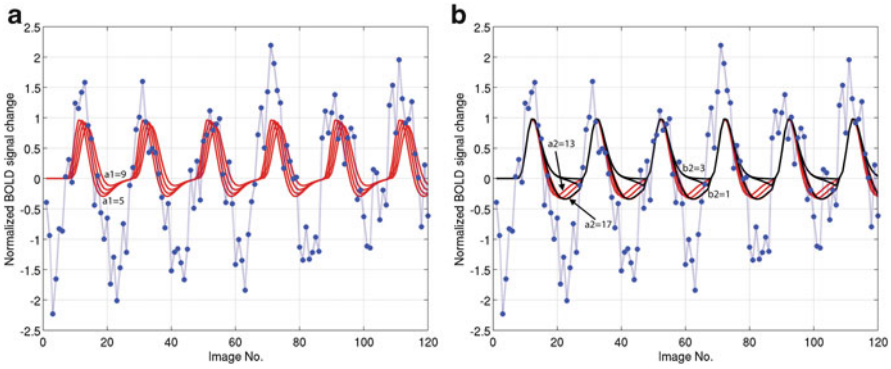

Fig. 2.10 Two-gamma function model for HRF. (**a**) The two-gamma function model for the response with varying parameter a_1; (**b**) The two-gamma function for the response with varying parameters a_2 and b_2

3. Obtain the angle and magnitude information corresponding to the fundamental frequency; this angle information is used to estimate the fMRI response delay.
4. Compute *ifft* (inverse FFT) using fundamental frequency magnitude information.
5. Adopt only the real value of the inverse FFT transformation to get the HRF fundamental frequency model, i.e., $f_{t,1}$ in Eq. (2.5).

Based on this algorithm, we obtain the estimated HRF model displayed as solid red curve in Fig. 2.7 (six cycles), Fig. 2.8b (ten cycles), and Fig. 2.9 (six cycles).

Another widely used function to model the BOLD-fMRI HRF is two-gamma function (Fig. 2.10a, b). To estimate the two-gamma function adaptively, we need to estimate the delay/onset of the function by the FFT analysis (Eq. (2.5)) or cross correlation analysis. Then, a two-gamma function is built according to the following equation [16]:

$$f_{t,2} = \left(\frac{t}{d_1}\right)^{a_1} \exp\left(-\frac{t-d_1}{b_1}\right) - c\left(\frac{t}{d_2}\right)^{a_2} \exp\left(-\frac{t-d_2}{b_2}\right) \tag{2.6}$$

where $a_1 = 6$, $b_1 = 0.9$, $d_1 = a_1 \times b_1$, $c = 0.35$, $b_2 = 0.9$, $a_2 = 12$, and $d_2 = a_2 \times b_2$ are the typical parameters. As shown in Fig. 2.10, the variation of these parameters will lead to poorer or better estimation of the HRF. In Fig. 2.10a, the parameter $a_1 = 5:2:10$; in Fig. 2.9b, the parameter $a_2 = 10:3:17$, while all the other parameters in Eq. (2.6) have not been changed.

Furthermore, the block/boxcar function (e.g., thin solid curve in Fig. 2.9) convolving with a Gaussian function can also be employed to model the HRF [17]. The function is expressed as

$$f_{t,3} = (\text{block}) \otimes \left(\exp\left(-\left(\frac{t}{\sqrt{2}c}\right)^2\right)\right) \tag{2.7}$$

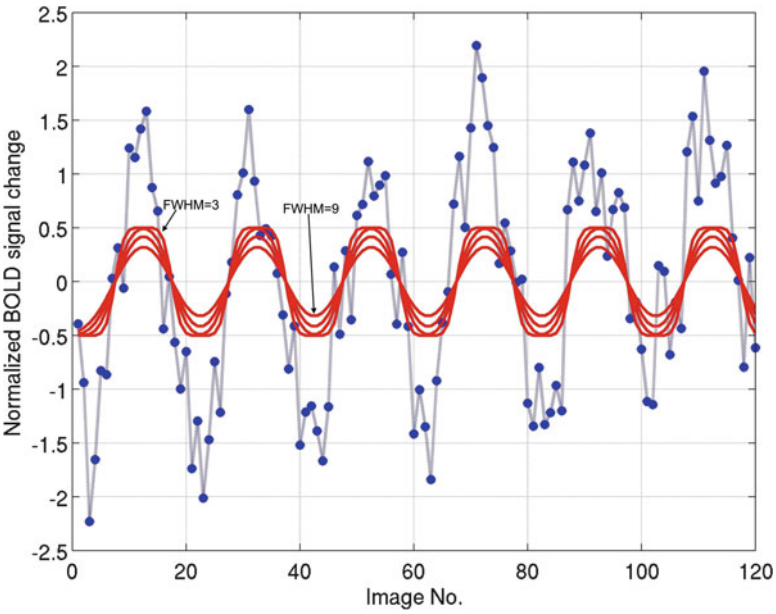

Fig. 2.11 One typical fMRI response and its models for the brain system input. Block function convolved with Gaussian function model response with varying parameter FWHM=3:2:9

where block is the block function and $\otimes$ is the convolution operation. The full width at half maximum (FWHM) is determined according to FWHM $= 2\sqrt{2\ \ln(2)}c$, and c is the standard deviation of the Gaussian function. Some of HRFs obtained from this method is displayed in Fig. 2.11. Comparing Fig. 2.11 with Fig. 2.9, we found that the HRF obtained from Gaussian convolution method is similar to the fundamental frequency HRF model if the wider FWHM is employed.

2.4.2 Drift Models for Activation Detection

Using GLM method for activation detection, we also need to model the slow drift in fMRI response. The low-frequency fMRI drift can be modeled by polynomial, spline, and cosine functions. In this study, we employ polynomial functions to model the fMRI response drift. For example, we show a maximum of 14th-order polynomials (i.e., $1, x, x^2, \ldots, x^{14}$) to model the drift. The shapes of the polynomials are given in Fig. 2.12. Figure 2.12a displays the curve of each polynomial against the time points, while the matrix format is given in Fig. 2.12b.

Low-frequency Fourier series (cosine/sine function) can also be used to model slow drift in fMRI time series; however, when using Fourier series to model the slow drift, the frequency range should be out of the range of stimuli because if we choose the drift frequency similar to the experimental design, it will be difficult to separate the signal (HRF) from the drift effect for activation detection.

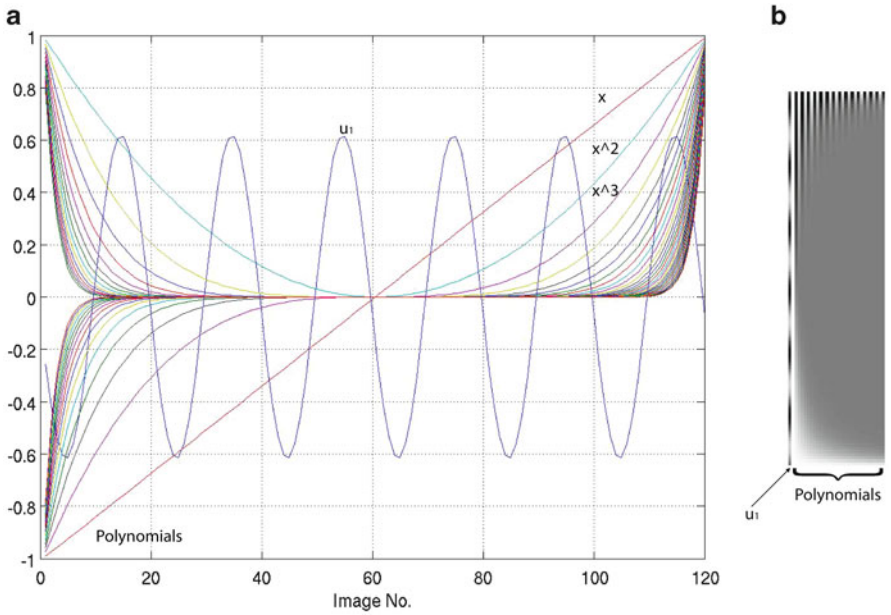

Fig. 2.12 Example of polynomial drifts in the model for the retinotopic mapping design with 120 fMRI image frames. (**a**) Curve for each polynomial term; (**b**) The corresponding matrix format. $u1$ is HRF obtained from FFT method

2.5 General Linear Model (GLM) for Activation Detection

After HRF model has been established, general linear model (GLM) for fMRI activation detection can be applied. It has been becoming one of the most commonly used methods since 1994 [18, 19]. The reason for its popularity is that the theory of linear model is relatively well developed in statistics. In addition, GLM offers advantages to combine and compare different runs/subjects within the framework of linear model. This is because the analysis of variance (ANOVA) is only a special case of linear model [20].

2.5.1 Generalized Linear Model (GLM) for Activation Detection

In the linear regression model for fMRI activation detection, we assume fMRI response y at each voxel within each run can be represented as

$$Y = X\beta + e \tag{2.8}$$

where $X = [X_1, X_2]$ is the design matrix, X_1 is the HRF model (Figs. 2.9 and 2.10), X_2 is the drift term which can be modeled by polynomials (e.g., Fig. 2.12), β is regression parameter that needs to be estimated, and $e \sim N(0,\sigma^2)$.

The model X in Eq. (2.8) can be determined beforehand or adaptively as shown in Sect. 2.4.1. The advantage of defining the model beforehand is that it greatly simplifies the analysis and interpretation of fMRI data if the model or the design matrix is selected correctly. It also simplifies the calculations if only a single fixed model for all neuron population responses in the brain is required. However, these targets are currently not feasible because the exact neuron response in different brain regions is unknown. More importantly, it is not reasonable to assume that all the neurons in the brain respond to the same stimulus in exactly the same way when it is presented on different occasions [16]. Also, inaccuracies will occur if one uses the same model (design matrix) for all the neuron populations in the brain. Therefore, it is necessary to build different models adaptively for different neuron populations in terms of their fMRI responses (e.g., Eq. (2.5) for phase-encoded design). This hypothesis is based on the fact that different neuron populations in the cortex have different responses (shape, magnitude, hemodynamic delay, slow drift, etc.) even for the same stimulus. To define the model adaptively using nonlinear regression method (i.e., Chap. 1) for each voxel is computational demanding; one easier way to achieve this goal is to choose the HRF from a predefined dictionary (e.g., models in Figs. 2.8, 2.9, and 2.10) which has the maximum correlation with the response. For the drift model, a fixed order of polynomials (i.e., the 5th order of polynomials seems to be enough for most study) is often employed, although model selection algorithm can also be used for drift model.

2.5.2 Ordinary Least Square for Parameter Estimation in GLM

Since both HRF and drift models have been established, the next step is to estimate the model parameters β in Eq. (2.8). One of most commonly used method for linear system parameter estimation is ordinary lest square method as shown in Chap. 1. To apply this method, we define the Euclidean distance between response and its estimation (fitting) results in the following quadratic form:

$$D = (Y - X\beta)^{\mathrm{T}} (Y - X\beta) \tag{2.9}$$

We want to get the system parameters β which minimize the distance D, i.e., minimize the sum square error (SSE) for the estimation. In this case, the least squares solution is equivalent to solving the following well-known normal equation:

$$\frac{\partial D}{\partial \beta} = 0 = -2X^{\mathrm{T}} (Y - X\beta) \tag{2.10}$$

From Eq. (2.9) and (2.10), it is easy to obtain estimation of β as [21]

$$\widehat{\beta} = X^{+}Y \tag{2.11}$$

where matrix $X^{+} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}$ is the pseudoinverse (PINV) of X and *pinv* in MATLAB can be used to estimate Moore–Penrose pseudoinverse of matrix method. Singular value decomposition (SVD) is one of the methods to compute matrix inverse (Eq. (1.65) in Chap. 1) which can be found in many books [22, 23]. In the next section, we introduce fast orthogonal search (FOS) algorithm which is based on Gram–Schmidt orthogonalization to calculate matrix inverse.

2.5.3 FOS to Solve the Inverse Problem

Fast orthogonal search algorithm [24] is similar to QR decomposition to estimate the matrix inverse. From Eq. (2.8), we have

$$Y = X\beta + e = \beta_0 x_0 + \beta_1 x_1 + \cdots + \beta_p x_p + e = \sum_{m=0}^{M} \beta_m P_m(t) + e \tag{2.12}$$

where $M = p, P_0(t) = x_0(t), P_1(t) = x_1(t), \ldots P_M(t) = x_p(t)$, and p is the total number of HRF and drift models to be included. Rearranging a sum of terms in Eq. (2.12) mutually orthogonal over the portion of data record extending from $t = N_0$ to $t = N$,

$$y(t) = \sum_{m=0}^{M} g_m W_m(t) + e \tag{2.13}$$

$y(t) = Y$ (where t is the image frame, N_0 is the first fMRI image frame, and N is the last image frame) and $W_m(t)$ is defined as

$$W_m(t) = P_m(t) - \sum_{r=0}^{m-1} \alpha_{mr} W_r(t) \tag{2.14}$$

where

$$\alpha_{mr} = \frac{\sum_{t=N_0}^{N} P_m(t) W_r(t)}{\sum_{t=N_0}^{N} (W_r(t))^2} \tag{2.15}$$

g_m in Eq. (2.13) is given by

$$g_m = \frac{\sum_{t=N_0}^{N} y(t)W_m(t)}{\sum_{t=N_0}^{N} (W_m(t))^2} \tag{2.16}$$

The coefficients g_m can be calculated as follows:

$$g_m = \frac{C(m)}{D(m,m)}, \quad m = 0, \ldots, M \tag{2.17}$$

where $D(0,0) = 1$, $D(m,0) = \overline{P_m(t)}, m = 1, \ldots, M$;

$$D(m,r) = \overline{P_m(t)P_r(t)} - \sum_{i=0}^{r-1} \alpha_{ri} D(m,i), \quad m = 1, \ldots, M; \ r = 1, \ldots, m. \tag{2.18}$$

$$\alpha_{mr} = \frac{D(m,r)}{D(r,r)}, \quad m = 1, \ldots, M; \ r = 0, \ldots, m-1 \tag{2.19}$$

with $C(0) = \overline{y(t)}$;

$$C(m) = \overline{y(t)P_m(t)} - \sum_{r=0}^{m-1} \alpha_{mr} C(r), \quad m = 1, \ldots, M. \tag{2.20}$$

Here we use the overbar (e.g., $\overline{P_m(t)P_r(t)}$) to denote the time average from $t = N_0$ to $t = N$. The coefficients β_m in Eq. (2.12) can be obtained in the following way:

$$\beta_m = \sum_{i=m}^{M} g_i V_i \tag{2.21}$$

$V_m = 1$ and

$$V_i = -\sum_{r=m}^{i-1} \alpha_{ir} V_r, \tag{2.22}$$

$i = m+1, \ldots, M$.

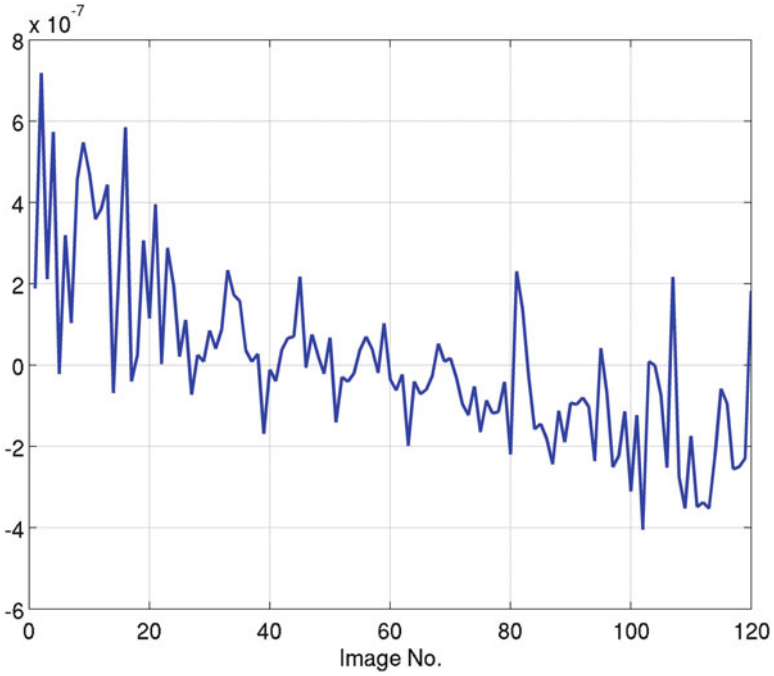

Fig. 2.13 Comparison between SVD and FOS methods to calculate matrix inverse

In combination with Fourier series, FOS algorithm has been applied to fMRI activation detection [25]. To show the performance of FOS algorithm for solving linear Eq. (2.8), we apply the algorithm to the fMRI response as shown in Fig. 2.8b. We found two methods, i.e., PINV and FOS produced very similar results. For the reason to compare it easily, we plot the difference between two algorithms, i.e., SVD and FOS methods in Fig. 2.13. The maximum estimated response difference is 8×10^{-7}. It is obvious that there is a very small difference between these two methods, and larger difference can be found in the time points where the magnitude of HRF is larger (compare Fig. 2.13 with Fig. 2.8b).

2.5.4 Weighted Least Squares Estimation

It should be noted that the estimation of β is based on the assumption that $e \sim N(0,\sigma^2)$ in Eq. (2.8). If the error term is $e \sim N(0, \sigma^2 W)$, where W is a positive definite known matrix, then weighted least squares method can be applied to estimate the system parameters. In the linear model (Eq. (2.8) with $E(y) = X\beta$), the generalized least square estimate of β is given by [26]

$$\widehat{\beta} = \left(X^{\mathrm{T}} W^{-1} X\right)^{-1} X^{\mathrm{T}} W^{-1} Y \tag{2.23}$$

This is because the distance we want to minimize is

$$D = (Y - X\beta)^{\mathrm{T}} W (Y - X\beta) \tag{2.24}$$

From Eq. (2.24), calculate $\frac{\partial D}{\partial \beta} = 0$, and we obtain Eq. (2.23).

Until now, we have not analyzed the error term e in Eq. (2.8) from an ordinary least squares fit. Exploratory analysis strongly suggests that the error term in the model for fMRI data analysis is in fact autocorrelated [27]. If we ignore this fact, we tend to get spurious declarations of significance when variables are really not important [13]. To remedy this problem, we need to model the autocorrelation error. One autocorrelated error structure frequently encountered in fMRI data analysis is the first-order autoregressive model (AR(1) model). Therefore, we start from AR(1) model, then introduce AR(q) (q is the order of autoregressive term) model for fMRI data analysis.

2.5.5 AR(1) Model

To remedy the autocorrelated error term in the GLM, the autocorrelation coefficient needs to be estimated. If the error term $e = e_t$ (t denotes the fMRI time frame) in Eq. (2.8) is the first order of autocorrelation [21, 27], then we model it as

$$e_t = \rho e_{t-1} + \varepsilon_t, \varepsilon_t \sim N\left(0, \tilde{\sigma}^2\right), \tag{2.25}$$

where $|\rho| < 1$ is the autocorrelation coefficient of the residuals. Let $y_e = \left[e_2\ e_3 \ldots e_n\right]^{\mathrm{T}}$ and $X_e = \left[e_1\ e_2 \ldots e_{n-1}\right]^{\mathrm{T}}$, then the least squares estimation of ρ is

$$\widehat{\rho} = \left(X_e' X_e\right)^{-1} X_e' y_e = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2}. \tag{2.26}$$

For the AR model, pre-whitening the data for each fMRI time series, i.e., pre-whitening Eq. (2.8), is now expressed as $Y_t = X_t\beta_t + e_t$, and consider its matrix form

$$\tilde{Y}_t = \tilde{X}_t \beta_t + \varepsilon_t \tag{2.27}$$

where

$$\tilde{Y}_1(1) = Y_1(1), \tilde{Y}_t = \frac{(Y_t - \rho Y_{t-1})}{\sqrt{1-\rho^2}}; \tag{2.28}$$

$$\tilde{X}_1(1) = X_1(1), \tilde{X}_t = \frac{(X_t - \rho X_{t-1})}{\sqrt{1-\rho^2}}, \tag{2.29}$$

$t = 2, \ldots, n$, where n is the total number of fMRI time frames. In the same way as Eq. (2.8), the estimation of β_t is

$$\widehat{\beta}_t = \tilde{X}^+ \tilde{Y}, \tag{2.30}$$

where $\tilde{X}^+ = \left(\tilde{X}^{\mathrm{T}}\tilde{X}\right)^{-1}\tilde{X}^{\mathrm{T}}$ is the pseudoinverse of $\tilde{X}$ which is the transformed design matrix and $\tilde{Y} = \left(\tilde{Y}_1, \tilde{Y}_2, \ldots, \tilde{Y}_n\right)'$ is the transformed observation vector (see Eqs. (2.29) and (2.28)). The vector of residuals $R = (r_1, \ldots, r_n)'$ is

$$R = \tilde{Y} - \tilde{X}\widehat{\beta}_t, \tag{2.31}$$

$$\widehat{\tilde{\sigma}}^2 = \frac{R'R}{v}, \tag{2.32}$$

where $v = n - \mathrm{rank}\left(\tilde{X}\right)$ is the degree of freedom (df).

2.5.6 AR(q) Model

If the error term follows the autoregressive (AR) process of order q, we call this model as AR(q) and it can be expressed as

$$e_t = \rho_1 e_{t-1} + \rho_2 e_{t-2} + \cdots + \rho_q e_{t-q} + \varepsilon_t, \tag{2.33}$$

where $\rho_1, \rho_2, \ldots,$ and ρ_q are parameters and ε_t has mean zero and satisfies the inference assumptions. To solve this problem, iterative method known as the Cochrane–Orcutt procedure can be employed [13]. It includes four major steps as follows:

1. Using Eq. (2.11) to get estimated β in Eq. (2.8), we get the corresponding residual e_t.
2. Applying least squares method to obtain parameters in Eq. (2.33), let $y_e = \left[\, e_2\ e_3\ \ldots\ e_n \,\right]^{\mathrm{T}}$ and $X_e = \left[\, e_1\ e_2\ \ldots\ e_{n-1} \,\right]^{\mathrm{T}}$, then the least squares estimation of AR parameters is $\rho = \left[\, \rho_1\ \rho_2\ \ldots\ \rho_q \,\right]^{\mathrm{T}} = \left(X_e' X_e\right)^{-1} X_e' y_e$.
3. Transferring (pre-whitening) Y and X in the following way:

$$\tilde{Y}_t = Y_t - \rho_1 Y_{t-1} - \rho_2 Y_{t-2} - \cdots - \rho_q Y_{t-q} \tag{2.34}$$

$t = q+1, q+2, \ldots, n$. Similarly, we need to transfer each column of the design matrix X. For instance, the constant term (zero order of polynomials) or

intercept column of n 1 s becomes a column of $n-q$ entries, each entry being $1-\rho_1-\rho_2-\cdots-\rho_q$. The other columns of the design matrix are transformed by replacing

$$\tilde{X}_{t,j} = X_{t,j} - \rho_1 X_{t-1,j} - \rho_2 X_{t-2,j} - \cdots - \rho_q X_{t-q,j} \tag{2.35}$$

then computing the new least squares point estimation (Eq. (2.27)) using the transformed data.

4. We use the new least squares point estimation and recomputed residuals and return to the second step. Then we iterate the process until least squares point estimates of the predictors do not change much between iterations. Usually, one or two iterations are sufficient.

2.6 Hypothesis Test and Threshold Correction

After all the parameters have been estimated, we can test the hypothesis if the brain activation introduced by the stimulus is significant or not. This becomes the problem of making inferences on the estimated parameters in the linear model as detailed below.

2.6.1 Hypothesis Testing for the Activation Detection

After β in GLM (Eq. (2.8) or β_t from GLM-AR method in Eq. (2.30)) has been estimated, the general linear hypothesis on β can be written as [26]

$$H_0 : C\beta = h \quad \text{versus} \quad H_1 : C\beta \neq h \tag{2.36}$$

where $C = [c_1, c_2, \ldots, c_m]$ is the contrast vector. $H_0 : C\beta = h$ is the null hypothesis and is expressed as a system of linear equations in β, and $H_1 : C\beta \neq h$ is the alternative hypothesis which means that at least one of the equations is not satisfied. Specially, we want to know whether there is an effect $C\beta \neq h$ or not, i.e., stimulus effect which is introduced by the external stimulus, i.e.,

$$c_1\beta_1 + c_2\beta_2 + \cdots + c_m\beta_m = h \tag{2.37}$$

In fMRI data analysis, we are mainly interested in the following two special cases. The first one is that we are interested in testing if a single model parameter is 0 or not, say, β_i is zero. The contrast vector C is a known row vector, $C = [0, \ldots, 0, 1, 0, \ldots, 0]$ where 1 corresponds to the HRF (system input) in the design matrix. For example, if β_1 is the coefficient associated with HRF at the first

column of the design matrix in GLM, then we can set $C = [1,0, \ldots,0]$. In this case, we want to test whether the brain activation introduced by the external stimulus is significant compared to the low-frequency drift or not. Thus, we consider the following hypothesis (set $h = 0$ in Eq. (2.36)):

$$H_0 : \beta_1 = 0 \quad \text{versus} \quad H_1 : \beta_1 \neq 0 \tag{2.38}$$

for single input (stimulus task) activation detection. The second case is the hypothesis that two of the parameters (corresponding to HRF for two different stimuli in the design matrix), say, β_1 and β_2, are equal against the two-side alternative that they differ. In this case, the contrast vector can be defined as $C = [1, -1, \ldots, 0]$. For example, we are interested in the response difference between a red and a blue visual stimulus, where 1 corresponds to the model of red stimulus and -1 corresponds to the model of the blue stimulus. After contrast vector C has been set, an effect can be defined as

$$E = C\widehat{\beta} \tag{2.39}$$

where $\widehat{\beta}$ is the estimated model parameter (β_t from Eq. (2.30) or β from Eq. (2.8)). The estimated standard deviation is

$$S = \|C\tilde{X}^+\|\widehat{\sigma} \tag{2.40}$$

where $\|\cdot\|$ is the matrix norm and $\widehat{\sigma}$ can be estimated as follows: define a vector of residuals $R = (r_1, \ldots,r_n)'$ as

$$R = Y - X\widehat{\beta}, \tag{2.41}$$

$$\widehat{\sigma}^2 = \frac{R'R}{v}, \tag{2.42}$$

where $v = n - \text{rank}\left(\tilde{X}\right)$ is the degree of freedom (df). Then the T statistic for the null hypothesis is

$$T = \frac{E}{S} \tag{2.43}$$

where T is often used to quantify the size of activation detected in the fMRI data. To detect more than one effect at the same time, that is, the rank of contrast matrix C is k, the T statistic is replaced by an F statistic defined as

$$F = \frac{E'\left(C\tilde{X}^+\left(C\tilde{X}^+\right)^{\mathrm{T}}\right)^{-1}E}{\left(k\widehat{\sigma}^2\right)} \tag{2.44}$$

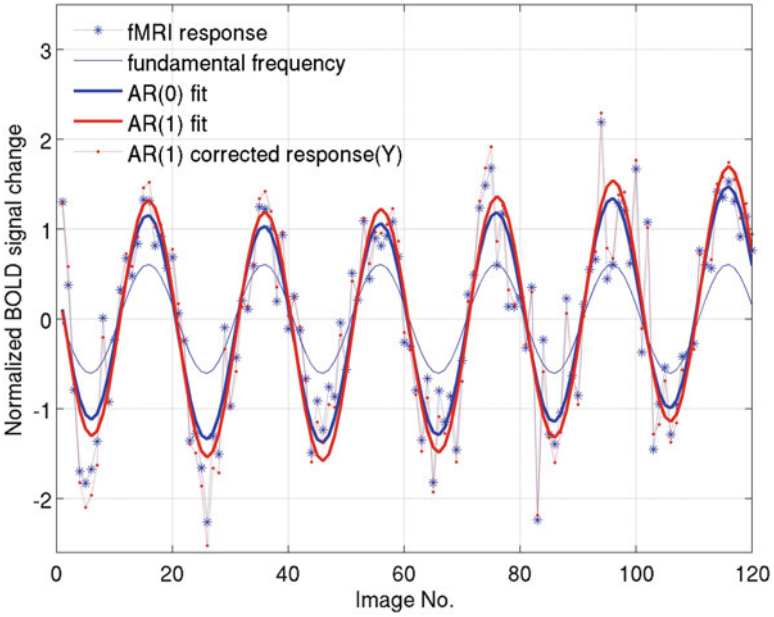

Fig. 2.14 Comparing AR(0) with AR(1) models

where $\tilde{X}$ is the transformed design matrix which can be obtained from Eq. (2.29). To demonstrate the T statistical method for activation detection, we have shown two examples using T value to quantify the brain activation in Figs. 2.3, and 2.8b. In these figures, we employ AR(1) model for the activation detection, i.e., with assuming that the error term in Eq. (2.8) is autocorrelated. In Fig. 2.3, the HRF was modeled by Gaussian function (FWHM = 5) convolution with boxcar function, the drift order is 8 (0–7 order of polynomials), and the T value is 13.7338. In Fig. 2.8b, we employed fundamental frequency model for HRF with the drift order of 6 (0–5 order of polynomials), and the T value we obtained is 11.7355. In both Figures, the thick solid curves represent the fitted responses using GLM method.

To compare the difference between AR model and non-AR model, we tested these algorithms for a single run activation detection from retinotopic mapping experiment from one subject with the right eye open. The stimulus was eccentricity normal order (expending ring stimulus) [28, 29]. The FFT fundamental frequency was used to model the system input, and 4 orders of polynomials were adopted to exclude the slow drift effects. One example of AR(0) and AR(1) fitting results is displayed in Fig. 2.14.

In Fig. 2.14, the dotted star blue curve represents the normalized fMRI response, while the red dotted dot curve denoted the pre-whiten data using the AR(1) model. The thin solid curve exhibits the HRF model using the fundamental frequency, the red think solid curve displays the AR(1) model fitted results, and the blue think solid

curve shows the fitted results without using AR algorithm. Based on the T statistics, we obtained $T = 18.4310$ using AR(0) model and $T = 21.4549$ using AR(1) model which is larger than the T value obtained from AR(0) model.

2.6.2 Bonferroni and FDR/FWE Threshold Correction

After T or F tests have been calculated for the activation detection, the final step is to perform multiple comparison threshold correction. Why do we need to do threshold correction in fMRI data analysis? To answer this question, let us take a simple example firstly. Assume we set the significant level as 5 % for brain activation detection and the brain size is 1,000 voxels, then we would have $1{,}000 \times 0.05 = 50$ voxels that will be active even if there is no task (false-positive). Therefore, we must control the false-positive in the results for activation detection. Bonferroni, false discover rate (FDR), and family-wise error rate (FWE) have been developed for multiple comparison in fMRI data analysis. We will introduce these methods for the final threshold correction in fMRI activation detection.

Bonferroni test [30] is one of the methods to achieve correction of spurious positives in the multiple comparison. To put it simple, let α denote the significance level and m represent the total number of (null) hypotheses testing, and then the Bonferroni criterion is

$$P \leq \frac{1}{m}\alpha \tag{2.45}$$

i.e., if the probability value is less than $\frac{1}{m}\alpha$, then the null hypotheses are true. To give an example to illustrate the method, we assume there are 15 comparisons in the brain region, and the ordered $P_{(i)}$s for the 15 comparisons made are 0.0001, 0.0004, 0.0019, 0.0095, 0.0201, 0.0278, 0.0298, 0.0344, 0.0459, 0.3240, 0.4262, 0.5719, 0.6528, and 0.7590,1.

These $P_{(i)}$ values can be obtained from T value in Eq. (2.43) with df, e.g., using *tinv* in MATLAB. In this example, if we set the significance level as $\alpha = 0.05$, then we have $P \leq 0.05/15 = 0.0033$ using the Bonferroni approach, and we reject the three hypotheses corresponding to the smallest P values. In fMRI data analysis, it means declare the voxel which P value is 0.0001, 0.0004, 0.0019 active, this activation is not by chance, or in other words, threshold the image of test statistics at the value corresponding to the P value $P \leq 0.05/15 = 0.0033$.

However, it has been criticized that this test is too conservative for multiple comparison. To overcome this limitation, false discover rate (FDR) and family-wise error rate (FWE) have been suggested to solve the false-positive problem [31, 32].

To demonstrate FDR and FWE for threshold correction, we consider testing $H_1, H_2, \ldots,$ and H_m based on the corresponding p values $P_1, P_2, \ldots,$ and P_m. Let $P_{(1)}, P_{(2)}, \ldots,$ and $P_{(m)}$ be the ordered p values and denote by $H_{(i)}$ the null hypothesis corresponding to $P_{(i)}$, where m is total number of (null) hypotheses testing

simultaneously. Define the following Bonferroni-type multiple-testing procedure, i.e., let k be the largest i for which

$$P_{(i)} \leq \frac{i}{m} q^*; \tag{2.46}$$

then reject all $H_{(i)}$ $i = 1, 2, \ldots, k$, where q^* is chosen to equal to the significance level α. In this study, we set $\alpha = 0.05$; m is the total number of comparison within the brain region mask (from Eq. (2.37)).

Using the same example as Bonferroni correction and the FDR controlling procedure with significant level $q^* = 0.05$, we now compare sequentially each $P_{(i)}$ with $P \leq 0.05i/15$, starting with $P_{(15)}$. The first P value to satisfy the constraint is

$$P_{(4)} = 0.0095 \leq 0.05 \times 4/15 = 0.0133$$

Thus, we reject the four hypotheses having P values which are less than or equal to 0.0133. Similarly, the FWE correction is using the following formula:

$$P_{(i)} \leq \frac{1}{m + 1 - i} q^* \tag{2.47}$$

In this example, we have FWE correction as

$$P_{(3)} = 0.0019 \leq 0.05/\,(15 + 1 - 3) = 0.0038$$

Therefore, we declare the first 3 voxels which have the smallest P value as active. In the example, Bonferroni and FWE declare the same voxel as active although the calculated threshold is different.

To illustrate these threshold correction methods for fMRI data analysis, we applied these algorithms to a single run activation detection from retinotopic mapping experiment from one subject with the right eye open. The stimulus was eccentricity normal order [28, 29] (see Appendix G for detail). The FFT fundamental frequency was used to model the system input with 4 order of polynomials to exclude the slow drift effects. The error term was modeled based on an AR(1) model as shown in Sect. 2.5.5. The contrast vector was defined as $c = [1,0,0,0,0]$, where 1 corresponds to the fundamental frequency to model the HRF (one example is shown in Fig. 2.14). A T statistical test (Eq. (2.43)) was employed to quantify the magnitude of activation. The activation was projected to the structural MRI of this subject in Talairach space [33] (Fig. 2.15). The X on the figure represents the x-axis coordinate of Talairach space. The color regions in Fig. 2.15 show where the brain regions are significantly activated by the stimuli (significant level at $P < 0.05$), and the color bar represents the corresponding magnitude. Figure 2.15a displays the activation threshold correction map using FDR method, while Fig. 2.15b shows FWE threshold correction results.

We used the visual cortex as a mask to get the total number of voxels m for the comparison. We estimated that there are 14,769 voxels (with voxel size 4 mm

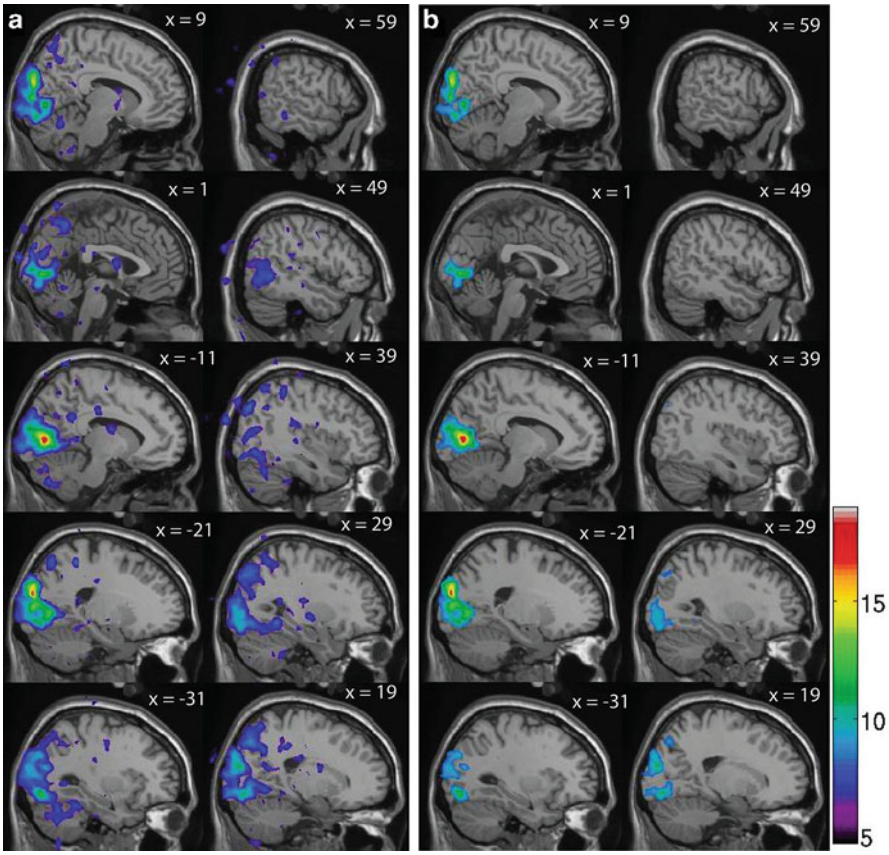

Fig. 2.15 Threshold correction for activation detection. (**a**) FDR method; (**b**) FWE method

by 4 mm by 4 mm) in the brain visual cortex region, and the estimated FDR threshold using Eq. (2.46) is 2.4916, FWE threshold is 4.7038 using Eq. (2.47), and Bonferroni threshold is 4.7174 based on Eq. (2.45). It can be seen that Bonferroni has the largest T value as threshold, suggesting it is the most conservative method for activation detection, followed by FWE algorithm. The activation map using Bonferroni correction method is similar to FWE correction method; therefore, we do not show the map here.

2.6.3 Number of Independent Tests

When we correct for multiple comparison using FDR or some other approaches, the adjusted α value (significant level) derives from the number of independent statistical tests. In fMRI analysis, this number could come from the total number

of voxels in the brain. However, in human brain, the adjacent voxels in HRF time series tend to correlate to each other, especially when the activation is derived by the large blood vessels. Considering the inter-voxel dependence in fMRI data, correction based on the number of voxels overestimates the number of independent spatial units, leading to too conservative α value. To determine a more accurate correction factor, techniques have been proposed that modify the denominator m in Bonferroni criterion (Eq. (2.45)) based on the degree of correlation between activated voxels [34].

2.6.4 Permutation/Random Test

The strength of Bonferroni and related methods are that they rely on only weak assumptions on dependence [35]. However, none of the methods make use of the spatial structure of the fMRI data. Random field theory and the resampling-based methods take into account for dependence in the fMRI data. Random resample method has becoming popular for fMRI data analysis due to the wide application of computer. Based on resample technique, permutation test [36] is a nonparametric method and is becoming more and more popular in statistical test. It has the striking advantage of distribution free, requires no assumption of an underlying correlation structure. However, this method may require large amounts of computer time to implement.

2.7 Summary of Algorithm for First-Level fMRI Data Analysis

fMRI study is complex and the general steps for data analysis include the following:

1. Design fMRI experiment according to the questions we want to address, and select type of experimental design, i.e., ER, block, or phase-encoded design. For data analysis we need to understand the experiment design for data processing using statistical methods (Sect. 2.1).
2. Data collection and preprocessing (e.g., motion correction and structural MRI image normalization etc.).
3. Normalize fMRI time series according to Eq. (2.1), and for model-free method, see Sects. 2.3.1 and 2.3.2. We focus on GLM method below.
4. Define HRF model for GLM using
   - Online approach (using Eq. (2.5) or provide several models for online selection), or
   - Offline approach (using Eq. (2.6) or (2.7) to define the model beforehand)

5. Select drift model for GLM analysis (Sect. 2.4.2).
6. Build GLM using Eq. (2.8) and estimate the model parameters β using SVD method (Eq. (2.11) or FOS method (Sect. 2.5.3)).
7. Correct the residual using AR(q) model as shown in Sects. 2.5.5 and 2.5.6.
8. Applying t-test for significant activation inference (Sect. 2.6.1).
9. Save E (effect map) from Eq. (2.39) and S (standard deviation map) from Eq. (2.40) for second-level analysis.
10. Threshold correction using Bonferroni-type method (Sect. 2.6.2).
11. Project the activation map to its structural MRI or a standard template for visualization.

Questions and Exercises

1. In the fMRI experimental design, why do we need to add control condition between each task stimulus? How long is the control condition interval between each task stimulus generally?
2. In an fMRI experiment, the subject for the study is often to be asked to respond to the stimuli by pressing the button. Why do we do this and how do we control the subject's attention in the scanner?
3. Many models have been used to model the hemodynamic response, for example, two-gamma-variate function, block function convolution with Gaussian function, and fundamental frequency of the block function. What is the advantage and disadvantage of these models for fMRI activation detection?
4. Compare two-sample t-test, correlation analysis method, and GLM methods for activation detection. What are the advantages and disadvantages for each method?
5. We have shown the linear method for activation detection; however, fMRI response is usually nonlinear; how do we use nonlinear regression method for fMRI response model? Assume the blood flow model is a gamma function for each block/task condition as shown in Chap. 1. What is the advantage and disadvantage of the nonlinear method?
6. In the phase-encoded design and standard block design, response delay can be estimated using FFT method. However, this method may not be suitable for ER and nonstandard block design; how do we use cross correlation method to estimate fMRI response delay for ER design?
7. In GLM, we need to model drift from fMRI response; a lower order of drift may lead to high T value, and a high order of drift model can result in a low T value, but it may overfit the response. How many polynomials should be included to model the slow drift in fMRI response for GLM?
8. Develop computer program to model the drift using spline function.
9. Design a contrast vector in GLM to compare red/green, blue/yellow, and achromatic, with control condition within the framework of GLM.
10. In the eccentricity stimuli from the retinotopic mapping experiments, we can estimate cell receptive field from the duty cycle of fMRI response. Develop computer program to estimate duty cycle.

11. The basic idea underlying neural adaptation and its study is that the neuronal system is plastic and repeated stimulation with the same set of stimuli results in automation and decreased activity in task-related regions. How to quantify adaptation effect in block design?

References

1. Boynton G, Engel SA, Glover GH, Heeger DJ (1999) Linear systems analysis of functional magnetic resonance imaging in human V1. J Neurosci 16(13):4207–4221
2. Vazquez AL, Noll DC (1998) Nonlinear aspects of the BOLD response in functional MRI. NeuroImage 7(2):108–118
3. Bandettini PA, Cox RW (2000) Event-related fMRI contrast when using constant interstimulus interval: theory and experiment. Magn Reson Med 43(4):540–548
4. Sereno M, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268:889–893
5. Engel S, Glover GH, Wandell BA (1997) Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb Cortex 7:181–192
6. Warnking J, Dojat M, Guérie-Dugué A, Delon-Martin C, Olympieff S, Richard N, Chehikian A, Segebarth C (2002) fMRI retinotopic mapping-step by step. NeuroImage 17(4):1665–1683
7. Hess RF, Li X, Mansouri B, Thompson B, Hansen BC (2009) Selectivity as well as sensitivity loss characterizes the cortical spatial frequency deficit in amblyopia. Hum Brain Mapp 30(12):4054–4069
8. Hess RF, Li X, Lu G, Thompson B, Hansen BC (2010) The contrast dependence of the cortical fMRI deficit in amblyopia: a selective loss at higher contrasts. Hum Brain Mapp 31:1233–1248
9. Derek LGH et al (2001) Medical image registration. Phys Med Biol 46(3):R1
10. Maintz JBA, Viergever MA (1998) A survey of medical image registration. Med Image Anal 2(1):1–36
11. Collins D, Neelin P, Peters TM, Evans AC (1994) Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J Comput Assist Tomogr 18:192–205
12. Hess RF, Thompson B, Gole G, Mullen KT (2009) Deficient responses from the lateral geniculate nucleus in humans with amblyopia. Eur J Neurosci 29:1064–1070
13. Sachs L (1984) Applied statistics: a handbook of techniques. Springer, New York
14. Bandettini PA et al (1993) Processing strategies for time-course data sets in functional MRI of the human brain. Magn Reson Med 30(2):161–173
15. Li X, Coyle D, Maguire L, McGinnity TM, Watson DR, Benali H (2010) A least angle regression method for fMRI activation detection for phase-encoded experimental designs. NeuroImage 52(2):1390–1400
16. Glover G (1999) Deconvolution of impulse response in event-related BOLD fMRI. NeuroImage 9(4):416–426
17. Rajapakse JC et al (1998) Modeling hemodynamic response for analysis of functional MRI time-series. Hum Brain Mapp 6(4):283–300
18. Friston KJ, Jezzard P, Turner R (1994) Analysis of functional MRI time-series. Hum Brain Mapp 1(2):153–171
19. Friston KJ et al (1995) Analysis of fMRI time-series revisited. NeuroImage 2(1):45–53
20. Seber G, Lee AJ (2003) Linear regression analysis, 2nd edn. Wiley, New York
21. Worsley K, Liao CH, Aston J, Petre V, Duncan GH, Morales F, Evans AC (2002) A general statistical analysis for fMRI data. NeuroImage 15:1–15

22. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) Numerical recipes: the art of scientific computing, 3rd edn. Cambridge University Press, Cambridge/New York
23. Golub G, van Loan C (1996) Matrix computations, 3rd edn, Johns Hopkins studies in mathematical sciences. The Johns Hopkins University Press, Baltimore
24. Korenberg M (1988) Identification nonlinear difference equation and functional expansion representations: the fast orthogonal algorithm. Ann Biomed Eng 16:123–142
25. Li X, Tian J, Wang X, Dia J, Ai L (2004) Fast orthogonal search method for modelling nonlinear hemodynamic response in fMRI. In: SPIE, Medical Imaging 2004: physiology, function, and structure from medical images, San Diego, CA, USA, 2004
26. Hocking RR (1996) Methods and applications of linear models, Wiley series in probability and statistics. Wiley, New York
27. Bullmore E, Brammer MJ, Williams SCR, Rabe-Hesketh S, Janot N, David AS, Mellers JDC, Howard R, Sham P (1996) Statistical methods of estimation and inference for functional MR images analysis. Magn Reson Med 35:261–277
28. Li X, Dumoulin SO, Mansouri B, Hess RF (2007) Cortical deficits in human amblyopia: their regional distribution and their relationship to the contrast detection deficit. Invest Ophthalmol Vis Sci 48:1575–1591
29. Li X, Dumoulin SO, Mansouri B, Hess RF (2007) The fidelity of the cortical retinotopic map in human amblyopia. Eur J Neurosci 25(5):1265–1277
30. Bland JM, Douglas GA (1995) Multiple significance tests: the Bonferroni method. BMJ 310(6973):170
31. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc B 57:289–300
32. Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29(4):1165–1188
33. Talairach J, Tournoux P (1998) Coplanar stereotaxic atlas of the human brain. Thieme, Stuttgart
34. Scott A, Huettel AWS, McCarthy G (2009) Functional magnetic resonance imaging, 2nd edn. Sinauer Associates Inc., Sunderland
35. Nichols T, Hayasaka S (2003) Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat Methods Med Res 12(5):419–446
36. Blair R, Karniski W (1993) An alternative method for significance testing of waveform difference potentials. Psychophysiology 30(5):518–524

http://www.springer.com/978-94-007-7302-8

Functional Magnetic Resonance Imaging Processing

Li, X.
2014, XIII, 221 p. 72 illus., 58 illus. in color.,
ISBN: 978-94-007-7302-8