

Abbreviations

DCM	Dynamic causal model
EPI	Echo planar imaging
fMRI	Functional magnetic resonance imaging
FFX	Fixed-effects analysis
FPR	False-positive rate
FWE	Family-wise error
FWHM	Full width at half maximum
GLM	General linear model
HRF	Haemodynamic response function
MIP	Maximum intensity projection
PET	Positron emission tomography
RFT	Random field theory
RFX	Random-effects analysis
SPM	Statistical parametric map(ping)
SVC	Small volume correction
TR	Time to repeat
VBM	Voxel-based morphometry

6.1 Introduction

Statistical parametric mapping (SPM) is an established statistical data analysis framework through which regionally specific effects in structural and functional neuroimaging data can be characterised. SPM is also the name of a free and open source academic software package through which this framework (amongst other things) can be implemented. In this chapter, we will give an overview of the underlying concepts of the SPM framework and will illustrate this by describing how to analyse a typical block-design functional MRI (fMRI) data set using the SPM software. An exhaustive description of SPM would be beyond the scope of this introductory chapter; for more information, we refer interested readers to *Statistical Parametric Mapping: The Analysis of Functional Brain Images* (Friston et al. 2007).

The aim of the SPM software¹ is to communicate and disseminate neuroimaging data analysis methods to the scientific community. It is developed by the SPM co-authors, who are associated with the Wellcome Trust Centre for Neuroimaging, including the Functional Imaging Laboratory, UCL Institute of Neurology. For those interested, a history of SPM can be found in a special issue of the *NeuroImage* journal, produced to mark 20 years of fMRI (Ashburner 2011). In brief, SPM was created by Karl Friston in approximately 1991 to carry out statistical analysis of positron emission tomography (PET) data. Since then, the SPM project has evolved to

G. Flandin (✉) • M.J.U. Novak
Wellcome Trust Centre for Neuroimaging,
UCL Institute of Neurology,
Queen Square, London WC1N 3BG, UK
e-mail: g.flandin@ucl.ac.uk; m.novak@ucl.ac.uk

¹<http://www.fil.ion.ucl.ac.uk/spm/>

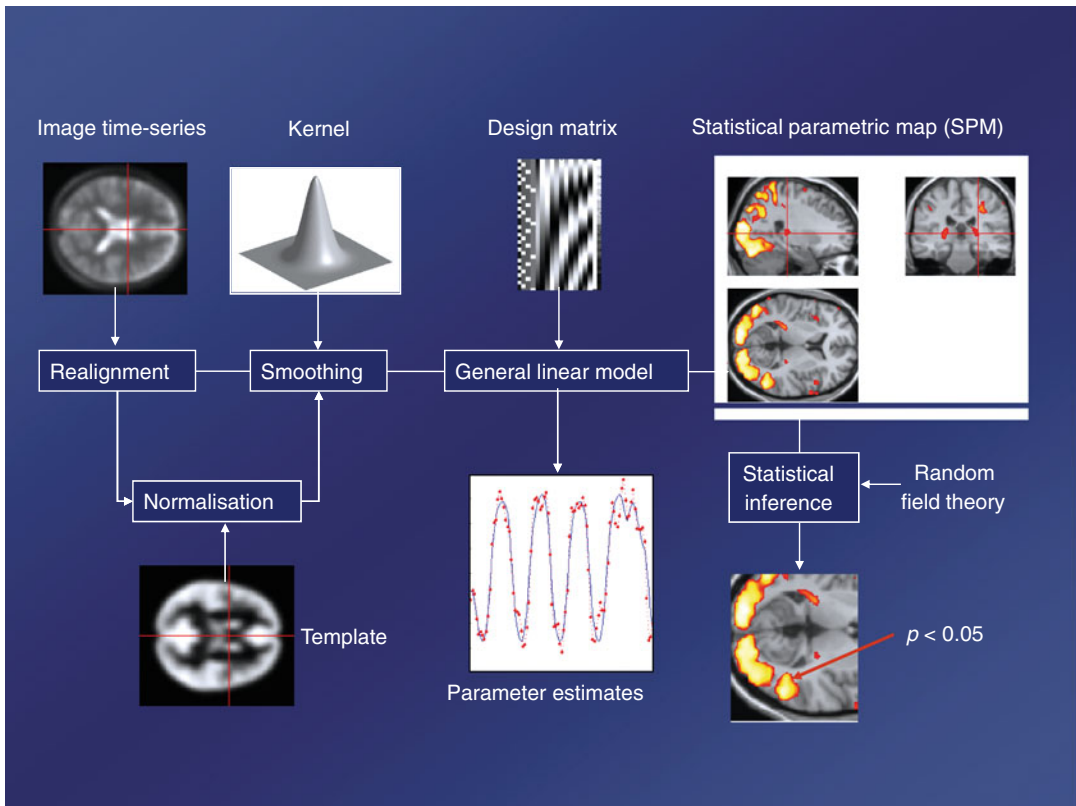


Fig. 6.1 Flowchart of the SPM processing pipeline, starting with raw imaging data and ending with a statistical parametric map (*SPM*). The raw images are motion-corrected, then subject to non-linear warping so that they match a template that conforms to a standard anatomical space. After smoothing, the general linear model is employed to estimate the parameters of a model encoded

by a design matrix containing all possible predictors of the data. These parameters are then used to derive univariate test statistics at every voxel; these constitute the SPM. Finally, statistical inference is performed by assigning p values to unexpected features of the SPM, such as high peaks or large clusters, through the use of the random field theory

support newer imaging modalities such as functional magnetic resonance imaging (fMRI) and to incorporate constant development and improvement of existing methods. The second half of the last decade saw an emphasis on the development of methods for the analysis of MEG and EEG (M/EEG) data, giving rise to the current version of SPM, SPM8.

The SPM framework is summarised in Fig. 6.1. The analysis pipeline starts with a raw imaging data sequence at the top left corner of the figure and ends with a statistical parametric map (also abbreviated to SPM) showing the significance of regional effects in the bottom right corner. The SPM framework can be partitioned into three key components, all of which will be described in this chapter:

- Preprocessing, or spatially transforming data: images are spatially aligned to each other to correct for the effect of subject movement during scanning (realignment/motion correction), then spatially normalised into a standard space and smoothed.
- Modelling the preprocessed data: parametric statistical models are applied at each voxel (a volume *element*, the three-dimensional extension of a pixel in 2D) of the data, using a general linear model (GLM) to describe the data in terms of experimental effects, confounding effects and residual variability.
- Statistical inference on the modelled data: classical statistical inference is used to test hypotheses that are expressed in terms of

GLM parameters. This results in an image in which the voxel values are statistics: this is a statistical parametric map (SPM). For such classical inferences, a multiple comparisons problem arises from the application of mass-univariate tests to images with many voxels. This is solved through the use of random field theory (RFT), resulting in inference based on corrected p values.

In this chapter, we will illustrate the concepts underpinning SPM through the analysis of an actual fMRI data set. The data set we will use was the first ever fMRI data set collected and analysed at the Functional Imaging Laboratory (by Geraint Rees, under the direction of Karl Friston) and is locally known as the Mother of All Experiments. The data set can be downloaded from the SPM website,² allowing readers to reproduce the analysis pipeline that we will describe on their own computers. (For more detailed step-by-step instructions to this analysis, we refer readers to the SPM manual.³) The purpose of the experiment was to explore equipment and techniques in the early days of fMRI. The experiment consisted of a single session in a single subject; the subject was presented with alternating blocks of rest and auditory stimulation, starting with a rest block. The auditory stimulation consisted of binaurally, bi-syllabic words presented at a rate of 60 words/min. Ninety-six whole brain echo planar imaging (EPI) scans were acquired on a modified 2 T Siemens MAGNETOM Vision System, with a repetition time (TR) of 7 s. Each block lasted for six scans, and there were 16 blocks in total, each lasting for 42 s. Each scan consisted of 64 contiguous slices ($64 \times 64 \times 64$, $3 \times 3 \times 3$ mm³ voxels). A structural scan was also acquired prior to the experiment ($256 \times 256 \times 54$, $1 \times 1 \times 3$ mm³ voxels).

Acquisition techniques have tremendously improved since this data set was acquired – a TR of 7 s seems very slow in comparison with today's standards – but the analysis pipeline is identical to that of more recent data sets and fits nicely with the purpose of illustration in this chapter.

While analysing this block, or epoch, designed experiment, we will point out the few steps that differ in the analysis of event-related data sets.

After an overview of the SPM software, we will describe in the next sections the three key components of an SPM analysis, namely, (i) spatial transformations, (ii) modelling and (iii) statistical inference.

6.2 SPM Software Overview

6.2.1 Requirements

The SPM software is written in MATLAB⁴ (The MathWorks, Inc.), a high-level technical computing language and interactive environment. SPM is distributed under the terms of the GNU General Public Licence. The software consists of a library of MATLAB M-files and a small number of C-files (which perform some of the most computationally intensive operations) and will run on any platform supported by MATLAB: 32- and 64-bit Microsoft Windows, Mac OS and Linux. This means that a prospective SPM user must first install commercially available software MATLAB. More specifically, SPM8 requires either MATLAB version 7.1 (R14SP3, released in 2005) or any more recent version (up to 7.13 (R2011b) at time of writing). Only core MATLAB software is required; no extra MATLAB toolboxes are needed.

A standalone version of SPM8, compiled using the MATLAB Compiler, is also available from the SPM development team upon request. This allows the use of most of the SPM functions without the need for a MATLAB licence (although this comes at the expense of being able to modify the software).

6.2.2 Installation

The installation of SPM simply consists of unpacking a ZIP archive from the SPM website on the user computer and then adding the root

²<http://www.fil.ion.ucl.ac.uk/spm/data/auditory/>

³<http://www.fil.ion.ucl.ac.uk/spm/doc/manual.pdf>

⁴<http://www.mathworks.com/>

SPM directory to the MATLAB path. If needed, more details on the installation can be found on the SPM wiki on Wikibooks.⁵

SPM updates (which include bug fixes and improvements to the software) take place regularly (approximately every 6 months) and are advertised on the SPM mailing list.⁶ SPM is a constantly evolving software package, and we therefore recommend that users either subscribe to the mailing list or check the SPM website so that they can benefit from ongoing developments. Updates can be easily installed by unpacking the update ZIP archive on top of the existing installation so that newer files overwrite existing ones. We would, however, advise users not to install updates mid-analysis (unless a specific update is needed), to ensure consistency within an analysis.

6.2.3 Interface

To start up SPM, simply type *spm* in the MATLAB command line and choose the modality in which you wish to use SPM in the new window that opens. A shortcut is to directly type *spm_fmri*. The SPM interface consists of three main windows, as shown in Fig. 6.2. The Menu window (1) contains entry points to the various functions contained in SPM. The Interactive window (2) is used either when SPM functions require additional information from the user or when an additional function-specific interface is available. The Graphics window (3) is the window in which results and figures are shown. Additional windows can appear, such as the Satellite Graphics window (4), in which extra results can be displayed, or the Batch Editor window (5). SPM can run in batch mode (in which several SPM functions can be set up to run consecutively through a single analysis pipeline), and the Batch Editor window is the dedicated interface for this. The window can be accessed through the ‘Batch’ button in the Menu window.

The Menu window is subdivided into three sections, which reflect the key components of an SPM analysis: spatial preprocessing, modelling and inference.

6.2.4 File Formats

In general, the first step when using SPM is to convert the raw data into a format that the software can read. Most MRI scanners produce image data that conform to the DICOM (Digital Imaging and Communications in Medicine) standard.⁷ The DICOM format is very flexible and powerful, but this comes at the expense of simplicity. As a consequence, the neuroimaging community agreed in 2004 to use a simpler image data format, the NIFTI (Neuroimaging Informatics Technology Initiative)⁸ format, to facilitate interoperability between fMRI data analysis software. The Mayo Clinic Analyze format was used prior to this but had several shortcomings which the NIFTI format overcame (including variability in the format versions used by different software packages which caused uncertainty about the laterality of the brain).

A NIFTI image file can consist either of two files, with the extensions *.hdr* and *.img*, or a single file, with the extension *.nii*. The two versions can be used in SPM interchangeably (note that you can also come across a compressed version of these files with a *.gz* extension; these are not supported in SPM and will need to be uncompressed outside the software before use). The *header (.hdr)* file contains meta-information about the data, such as the voxel size, the number of voxels in each direction and the data type used to store values. The *image (.img)* file contains the raw 3D array of voxel values. A file with the *.nii* extension contains all of this information in one file. A key piece of information stored in the header is the *voxel-to-world mapping*: this is a spatial transformation that maps from the stored data coordinates (voxel column *i*, row *j*, slice *k*) into a real-world

⁵<http://en.wikibooks.org/wiki/SPM/>

⁶<http://www.fil.ion.ucl.ac.uk/spm/support/>

⁷<http://dicom.nema.org/>

⁸<http://nifti.nimh.nih.gov/>

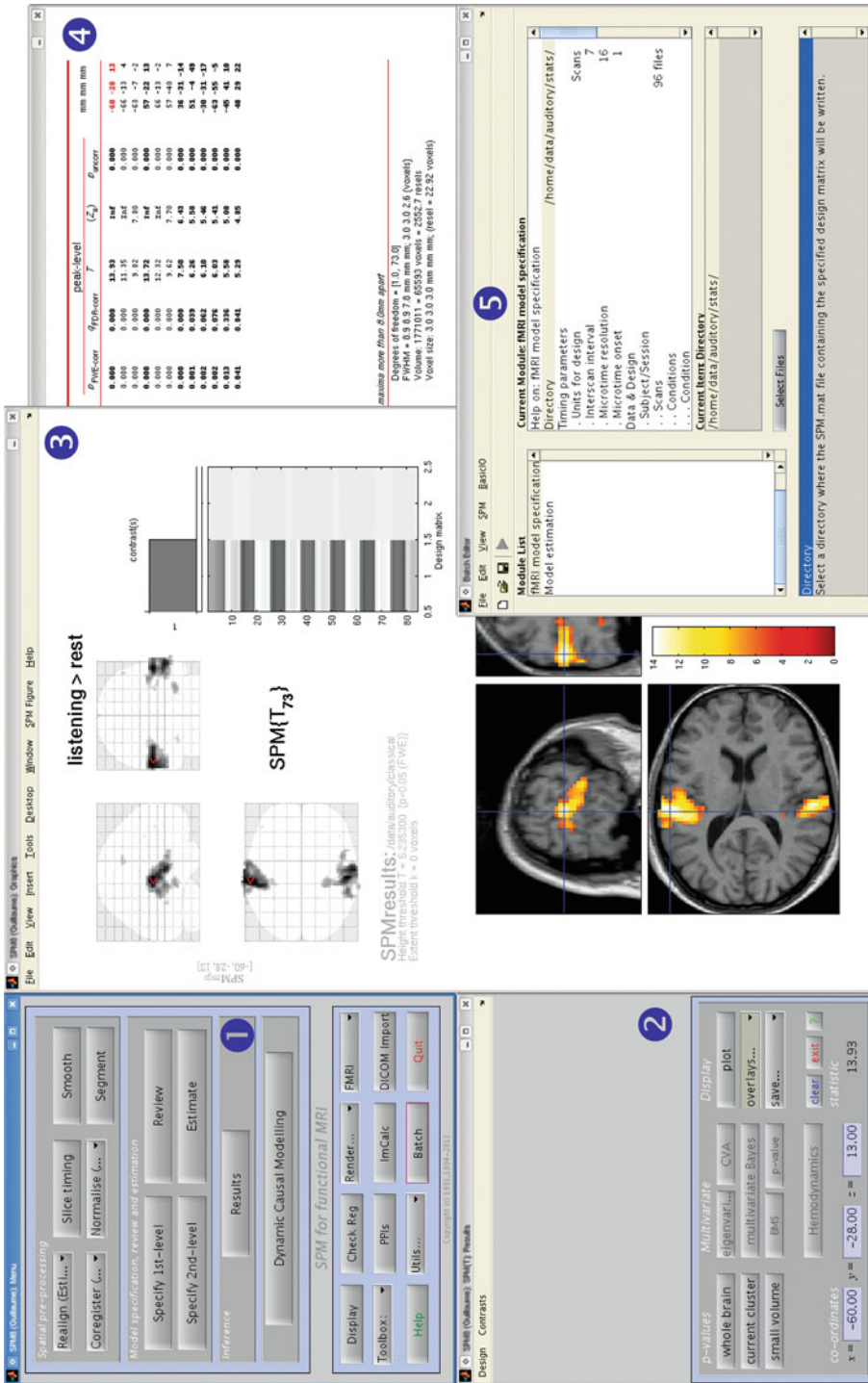


Fig. 6.2 Interface of the SPM8 software for the fMRI modality: (1) Menu window, (2) Interactive window, (3) Graphics window, (4) Satellite Results window and (5) Batch editor interface

position (x, y, z mm) in space. The real-world position can be in either a standardised space such as Talairach and Tournoux space or Montreal Neurological Institute (MNI) space or a subject-specific space based on scanner coordinates.

fMRI data can be considered as 4D data – a time series of 3D data – and can therefore be stored as a single file in the NIfTI format where the first three dimensions are in space and the fourth is in time. However, use of multiple 3D (spatial) files rather than a single 4D file is recommended with SPM for the time being because the software handles them more efficiently.

DICOM image data can be converted into NIfTI files in SPM using the ‘DICOM Import’ button in the Menu window. This is usually a straightforward process. If needed, however, NIfTI data obtained from the file converter of any other software package (such as the LONI Debabeler⁹ or dcm2nii¹⁰) can also be used in SPM; the output NIfTI images are interoperable between software packages.

The auditory data set used in this chapter has already been converted from the DICOM format. We have 96 pairs of fMRI files, namely *fM00223_*. {hdr,img}*, and one pair of structural files, namely *sM00223_002. {hdr,img}*. These images are actually stored in the Analyze format because they were acquired prior to 2004; SPM8 can read Analyze format as well as NIfTI format but will save new images in the NIfTI format.

Images can be displayed in SPM using the ‘Display’ and ‘Check Reg’ buttons from the Menu window. The first function displays a single image and some information from its header, while the second displays up to 15 images at the same time. This can be used to check the accuracy of alignment, for example.

6.3 Spatial Transformations

A number of preprocessing steps must be applied to the fMRI data to transform them into a form suitable for statistical analysis. Most of these

steps correspond to some form of image registration, and Fig. 6.3 illustrates a typical preprocessing pipeline. There is no universal pipeline to use in all circumstances – options depend on the data themselves and the aim of the analysis – but the one presented here is fairly standard.

The first preprocessing step is to apply a motion correction algorithm to the fMRI data (this is the *realignment function*). This may eventually include some form of distortion correction. A structural MRI of the same subject is often acquired and should be brought into alignment with the fMRI data in a second step (*coregister function*). The warps needed to spatially normalise the structural image into some standard space should then be estimated (*normalise function*) and applied to the motion-corrected functional images to normalise them into the same standard space (*write normalise function*). The final step will typically be to smooth the functional data spatially by applying a Gaussian kernel to them (*smooth function*).

The type of spatial *transformations* that should be applied to data depends on whether the data to be transformed all come from the same subject (*within-subject* transformations) or from multiple different subjects (*between-subject* transformations). The choice of *objective function* (the criterion to assess the quality of the registration) used to estimate the deformation also depends on the modality of data in question. Realignment is a *within-subject, within-modality* registration, while coregistration is a *within-subject, between-modality* registration. Normalisation is a *between-subject* registration. Within-subject registration will generally involve a *rigid body* transformation, while a between-subject registration will need estimation of affine or non-linear warps; this is because a more complex transformation is required to warp together the anatomically variable brains of different subjects than to warp together different images of the same brain. A criterion to compare two images of the same modality can be the sum of squares of the differences of the two images, while the comparison between two images of different modalities will involve more advanced criteria. An *optimisation algorithm* is then used in the registration step to

⁹<http://loni.ucla.edu/Software/Debabeler>

¹⁰<http://www.cabiatl.com/mricro/mricron/dcm2nii.html>

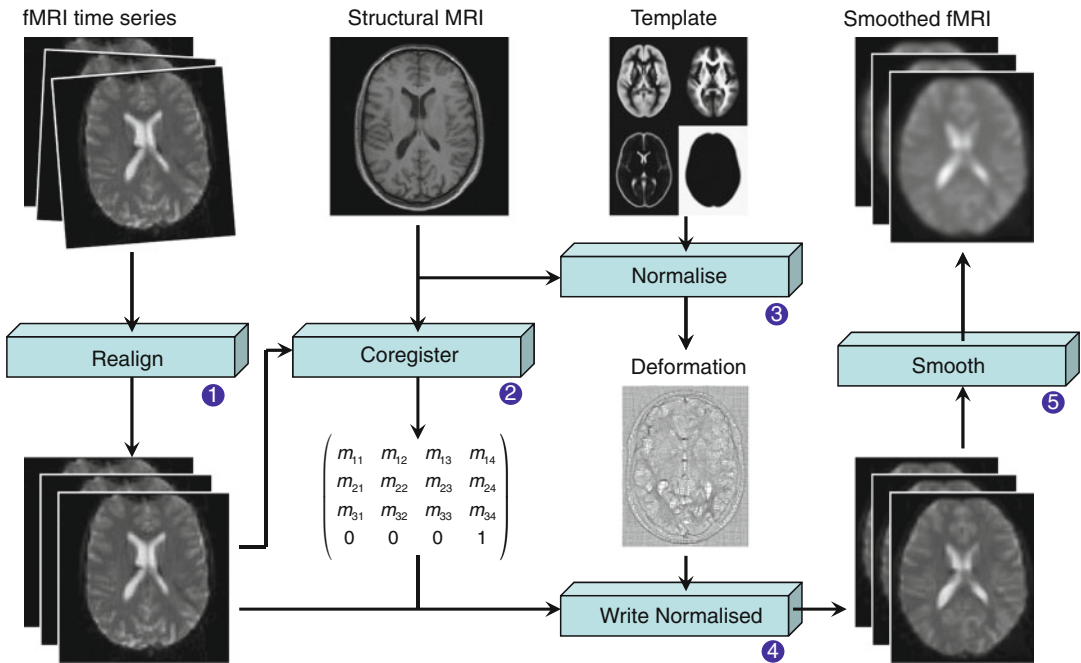


Fig. 6.3 Flowchart of a standard pipeline to preprocess fMRI data. After realignment (1) to correct for movement, structural and functional images are coregistered (2) then

normalised (3, 4) to conform to a standard anatomical space (e.g. MNI space) before being spatially smoothed using a Gaussian kernel (5)

maximise (or minimise) the objective function. Once the parameters have been estimated, the target image can be transformed to match the source image by resampling the data using an *interpolation* scheme. This step is referred to as *reslicing* when dealing with rigid transformations.

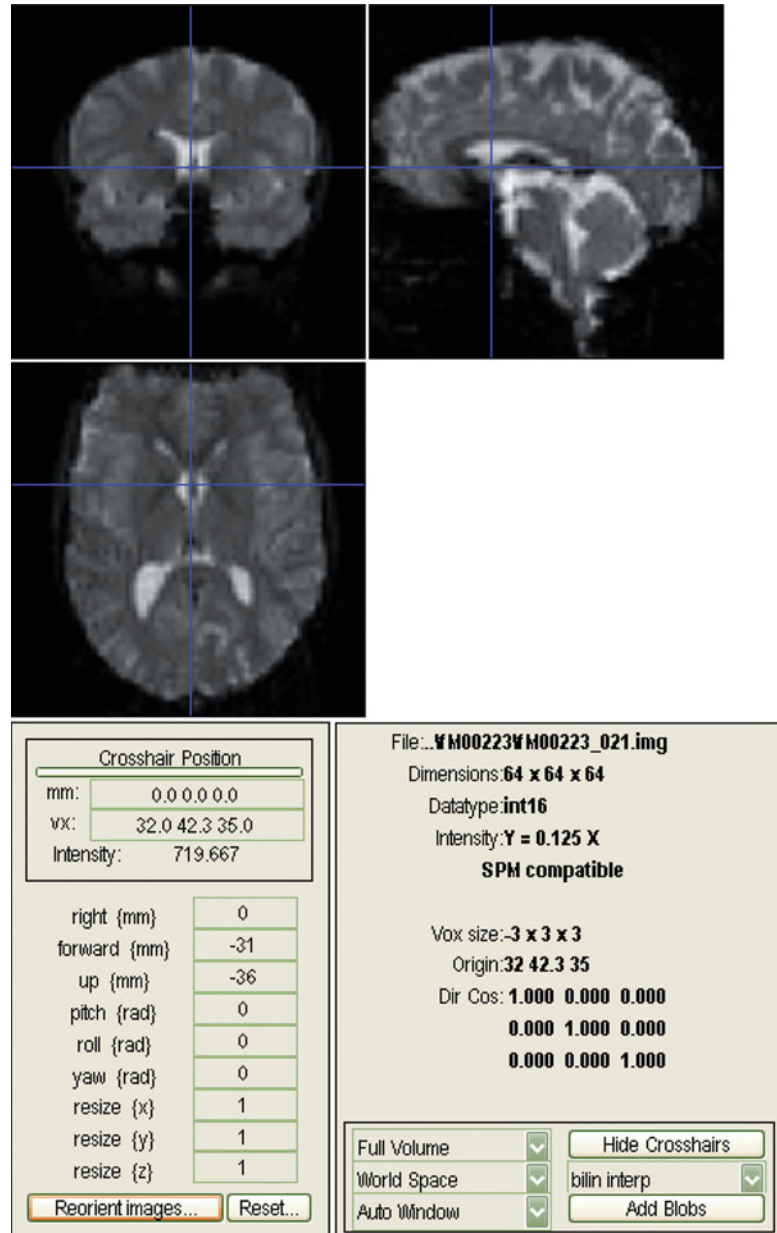
6.3.1 Data Preparation

Before preprocessing an fMRI data set, the first images acquired in a session should be discarded. This is because much of the very large signal change that they contain is due to the time it takes for magnetisation to reach equilibrium. This can be easily seen by looking at the first few images at the beginning of the time series using the display function. Some scanners might handle these ‘dummy scans’ automatically by acquiring a few scans before the real start of the acquisition, but this should be checked. In our example data set, we are going to discard the first 12 scans, leaving us with 84 scans. This is more than necessary here but it preserves the simplicity of the

experimental design as it corresponds to one complete cycle of auditory stimulation and rest.

It is good practice to manually reorient the images next so that they roughly match the normalised space that SPM uses (MNI space). This will help the convergence of the registration algorithms used in preprocessing; the algorithms use a local optimisation procedure and can fail if the initial images are not in rough alignment. In practice, the origin (0, 0, 0 mm) should be within 5 cm of the anterior commissure (a white matter track which connects the two hemispheres across the midline), and the orientation of the images should be within about 20° of the SPM template. To check the orientation of images, display one image of the time series using the ‘Display’ button and manually adjust their orientation using the translation (right, forward, up) and rotation (pitch, roll, yaw) parameters in the bottom left panel until the prerequisites are met. To actually apply the transformation to the data, you need to press the ‘Reorient images’ button and select all the images to reorient. With the auditory data set, the structural image is already correctly orientated,

Fig. 6.4 Interface of the ‘Display’ option. The location of the crosshair, in blue, is indicated in the bottom left panel, both in mm and voxel. Here, a translation of $[0, -31, -36]$ mm allows to set the origin of the image $([0, 0, 0]$ mm) near the anterior commissure



but the functional scans should be translated by about $[0, -31, -36]$ mm. See Fig. 6.4 for a screenshot of the ‘Display’ interface illustrating how to change the origin of a series of images.

6.3.2 Realignment

As described above, the first preprocessing step is to realign the data to correct for the effects of subject movement during the scanning session.

Despite restraints on head movement, cooperative subjects still show displacements of up to several millimetres, and these can have a large impact on the significance of the ensuing inference; in the unfortunate situation where a subject’s movements are correlated with the experimental task, spurious activations can be observed if no correction was performed prior to statistical analysis. Alternatively, movements correlated with responses to an experimental task can inflate unwanted variance components

in the voxel time series and reduce statistical power.

The objective of realignment is to determine the rigid body transformation that best maps the series of functional image volumes into a common space. A rigid body transformation can be parameterised by six parameters in 3D: three translations and three rotations. The realignment process involves the estimation of the six parameters that minimise the mean squared difference between each successive scan and a reference scan (usually the first or the average of all scans in the time series) (Friston et al. 1995).

Unfortunately, even after realignment, there may still be some motion-related artefacts remaining in the functional data (Friston et al. 1996b); this is mainly because the linear, rigid body realignment transformation cannot capture non-linear effects. These non-linear effects can be the consequence of subject movement between slice acquisition, interpolation artefacts, magnetic field inhomogeneities or spin-excitation history effects. One solution is to use the movement parameter estimates as covariates of no interest during the modelling of the data. This will effectively remove any signal that is correlated with functions of the movement parameters but can still be problematic if the movement effects are correlated with the experimental design. An alternative option is to use the ‘Realign and Unwarp’ function (Andersson et al. 2001). The assumption in this function is that the residual movement variance can be largely explained by susceptibility-by-movement interaction: the non-uniformity of the magnetic field is the source of geometric distortions during magnetic resonance acquisition, and the amount of distortion depends partly on the position of the head of the subject within the magnetic field; hence, large movements will result in changes in the shape of the brain in the images which cannot be captured by a rigid body transformation. The ‘Realign and Unwarp’ function uses a generative model that combines a model of geometric distortions and a model of subject motion to correct images. The ‘Realign and Unwarp’ function can be combined with the use of field maps (see the FieldMap toolbox), to further correct those geometric distortions introduced during the echo planar imaging (EPI) acquisition

(Jezzard and Balaban 1995). The resulting corrected images will have less movement-related residual variance and better matching between functional and structural images than will the uncorrected images.

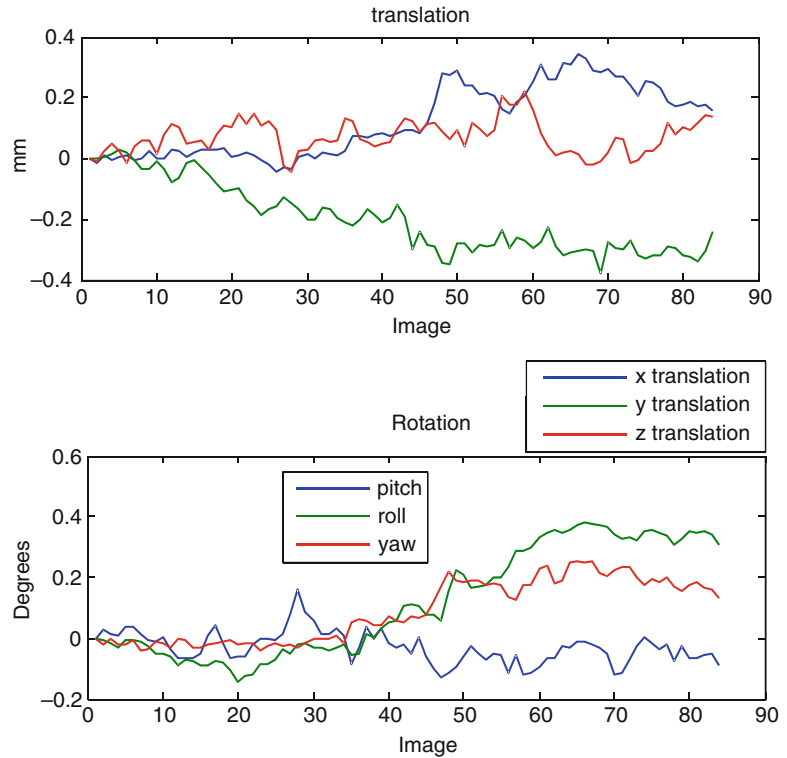
For the auditory data set, functional data are motion-corrected using ‘Realign: Estimate and Reslice’. Data have to be entered session by session to account for large subject movements between sessions that the algorithm is not expecting. With regard to the reslicing, it is sufficient to write out only the mean image. The output of the estimation will be encoded in the header of each image through modification of the original voxel-to-world mapping. It is best to reslice the data just once at the end of all the preprocessing steps; this ensures that all the affine transformations are taken into account in one step, preventing unnecessary interpolation of the data. The estimated movement parameters (see Fig. 6.5) are saved in a text file in the same folder as the data with an ‘*rp_*’ prefix and will be used later in the analysis.

6.3.3 Coregistration

Coregistration is the process of registering two images of the same or different modalities from the same subject; the intensity pattern might differ between the two images, but the overall shape remains constant. Coregistration of single subject structural and functional data firstly allows functional results to be superimposed on an anatomical scan for clear visualisation. Secondly, spatial normalisation is more precise when warps are estimated from a detailed anatomical image than from functional images; if the functional and structural images are in alignment, warps estimated from the structural image can then be applied to the functional data.

As with realignment, coregistration is performed by optimising the six parameters of a rigid body transformation; however, the objective function is different as image intensities cannot be compared directly as they were with the sum of squared differences. Instead, the similarity measures that are used rely on a branch of applied mathematics called *information theory* (Collignon et al. 1995; Wells et al. 1996). The

Fig. 6.5 Plot of the estimated movement parameters (three translations and three rotations) for the auditory data set as a function of time (or scan). Movements are small for this subject but we can still observe some slow drifts over time



most commonly used similarity measure is called *mutual information*; this is a measurement of shared information between data sets, based on joint probability distributions of the intensities of the images. The mutual information is assumed to be maximal when the two images are perfectly aligned and will serve here as the objective function to maximise.

For the auditory data set, the structural image should be coregistered to the mean functional image (computed during realignment) using ‘Coregister: Estimate’. Once again, there is no need to reslice at this stage; reslicing can be postponed until later to minimise interpolation steps. In the interface, the *reference* image (the target) is the mean functional image, while the *source* image is the structural image. Default parameters can be left as they are; they have been optimised over years and should satisfy most situations. The output of the algorithm will be stored in the header of the structural image by adjusting its voxel-to-world mapping. Figure 6.6 shows the alignment of the two images after registration.

6.3.4 Spatial Normalisation

Spatial normalisation is the process of warping images from a number of individuals into a common space. This allows signals to be compared and averaged across subjects so that common activation patterns can be identified: the goal of most functional imaging studies. Even single subject analyses usually proceed in a standard anatomical space so that regionally specific effects can be reported within a frame of reference that can be related to other studies (Fox 1995). The most commonly used coordinate system within the brain imaging community is the one described by Talairach and Tournoux (1988), although new standards based on digital atlases (such as the Montreal Neurological Institute (MNI) space) are nowadays widespread (Mazziotta et al. 1995).

The rigid body approach used previously when registering brain images from the same subject is not appropriate for matching brain images from different subjects; it is insufficiently complex to deal with interindividual differences

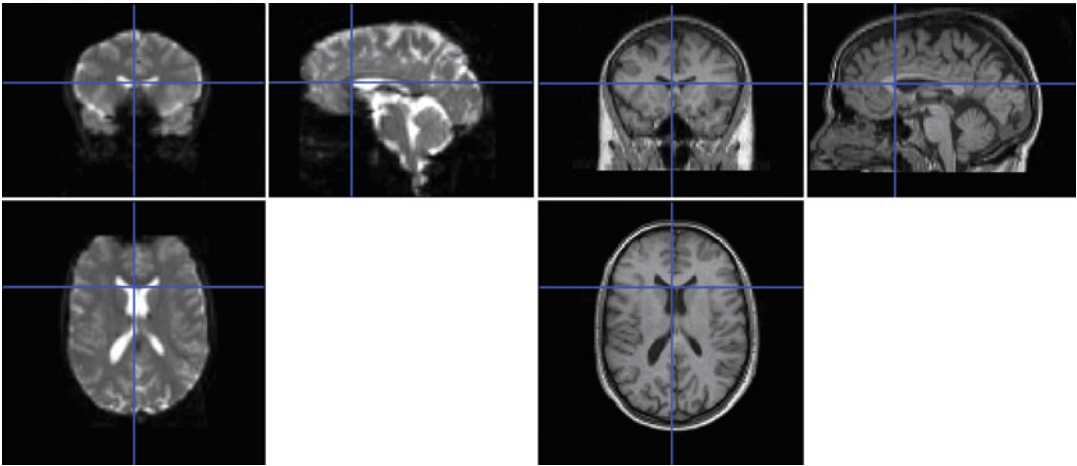


Fig. 6.6 Coregistered mean functional and structural images using mutual information. Some important dropouts in the functional data can be observed

in anatomy. More complex transformations (i.e. with more degrees of freedom) such as affine or *non-linear transformations* are used instead. (Non-linear registration is also used when characterising change in a subject's brain anatomy over time, such as those due to growth, ageing, disease or surgical intervention.)

The normalisation deformation model has to be flexible enough to capture most changes in shape but must also be sufficiently constrained that realistic brain warps are generated; *a priori*, we expect the deformation to be spatially smooth. This can be nicely framed in a Bayesian setting by adding a prior term to the objective function to incorporate prior information or add constraints to the warp. For instance, consider a deformation model in which each voxel is allowed to move independently in three dimensions. There would be three times as many parameters in this model than there are voxels. To deal with this, the deformation parameters need to be regularised; the prior term enables this. Priors become more important as the number of parameters specifying the mapping increases, and they are central to high-dimensional non-linear warping schemes. The approach taken in SPM is to parameterise the deformations by a linear combination of smooth, continuous basis functions, such as low-frequency cosine transform basis functions (see Fig. 6.7) (Ashburner and Friston 1999). These models

have a relatively small number of parameters, about 1,000 (although this is of course large in comparison with 6 parameters for a rigid body transformation and 12 for an affine transformation), and allow a better description of the observed structural changes whilst providing reasonably smoothed deformations. The optimisation procedure relies on an iterative local optimisation algorithm and needs reasonable starting estimates (hence the reorientation of the images at the very beginning of the analysis pipeline). This is the model underlying the 'Normalise' button in SPM. For this function, the user should select a template image (in MNI space) in the same modality as the experimental image to be normalised.

In practice, better alignment can be achieved by matching grey matter with grey matter and white matter with white matter. The process of classifying voxels into different tissue types is called *segmentation*, and an approach combining segmentation and normalisation will provide better results than normalisation alone. This is the strategy implemented via the 'Segment' button in SPM; it is referred to as *Unified segmentation* (Ashburner and Friston 2005). Unified segmentation uses a generative model which involves (i) a mixture of Gaussians to model intensity distributions, (ii) a bias correction component to model smooth intensity variations in

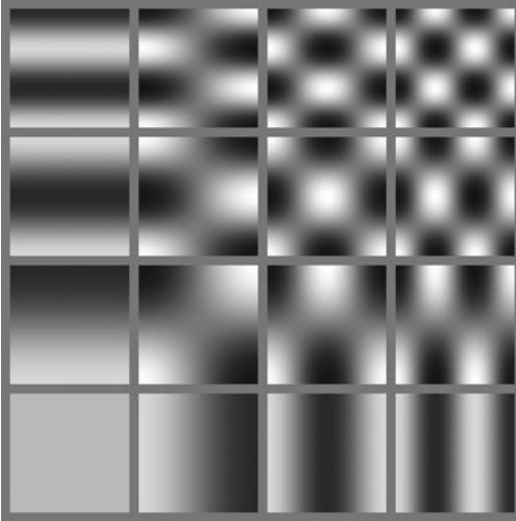


Fig. 6.7 Cosine transformation basis functions (in 2D) used by normalisation and unified segmentation. They allow deformations to be modelled with a relatively small number of parameters. The basis function registration estimates the global shapes of the brains, but is not able to account for high spatial frequency warps

space and (iii) a non-linear registration with tissue probability maps, parameterised using the low-dimension approach described in the previous paragraph.

We will use the unified segmentation approach on the auditory data set. The segmentation of the structural image (using, by default, tissue probability maps of grey matter, white matter and cerebrospinal fluid that can be found in the ‘*tpm*’ folder of the SPM installation) will generate a few files: images with prefixes ‘*c1*’ and ‘*c2*’ are estimated maps of grey and white matter, respectively, while the image with an ‘*m*’ prefix is the bias-corrected version of the structural image. Importantly, the estimated parameters of the deformation are saved in a MATLAB file ending with ‘*seg_sn.mat*’. This file can be used to apply the deformation, that is, to normalise the functional images (as they are in the same space as the structural scan thanks to the coregistration step) with the ‘Normalise: Write’ button. A new set of 84 images will be written to disk with a ‘*w*’ prefix (for warp). The same procedure can be applied to the bias-corrected structural scan in order to later superimpose the functional activations on the

anatomy of the subject. In both instances, some parameters have to be updated: the voxel size of the new set of images is preferably chosen in relation to the initial resolution of the images, for example, to the nearest integer. Here we used [3, 3, 3] for the functional data and [1, 1, 3] for the structural scan. Also, the interpolation scheme can be changed to use higher-order interpolation (the default is trilinear), such as a fourth-degree B-spline (Unser et al. 1993). The coordinates of locations within a normalised brain can now be reported as MNI coordinates in publications (Brett et al. 2002).

On a final note, spatial normalisation may require some extra care when dealing with patient populations with gross anatomical pathology, such as stroke lesions. This can generate a bias in the normalisation as the generative model is based on anatomically ‘normal’ data. Solving this usually involves imposing constraints on the warping to ensure that the pathology does not bias the deformation of undamaged tissue, for example, by decreasing the precision of the data in the region of pathology so that more importance is afforded to the anatomically normal priors. This is the principle of lesion masking (Brett et al. 2001). There is evidence, however, that the Unified Segmentation approach is actually quite robust in the presence of focal lesions (Crinion et al. 2007; Andersen et al. 2010).

6.3.5 Spatial Smoothing

Spatial smoothing consists of applying a spatial low-pass filter to the data. Typically, this takes the form of a 3D Gaussian kernel, parameterised by its full width at half maximum (FWHM) along the three directions. In other words, the intensity at each voxel is replaced by a weighted average of itself and its neighbouring voxels, where the weights follow a Gaussian shape centred on the given voxel. The underlying mathematical operation is a convolution, and the effect of smoothing with different kernel sizes is illustrated in Fig. 6.8.

It might seem counterintuitive to reduce the resolution of fMRI data through smoothing, but there are four reasons for doing this. Firstly,

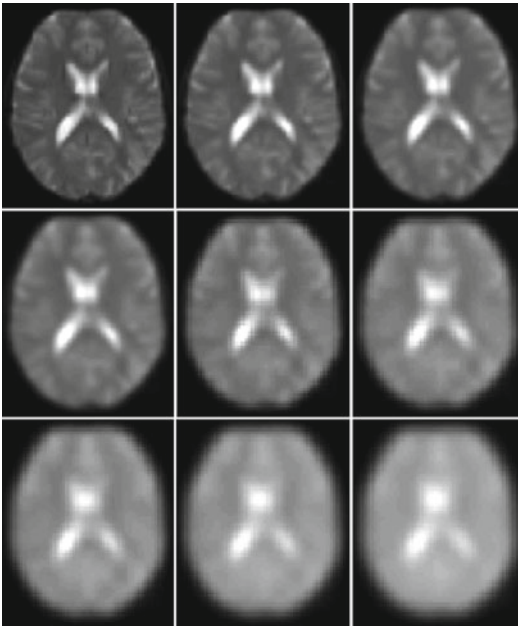


Fig. 6.8 Axial slice of a functional MRI scan smoothed with 3D Gaussian kernels of different isotropic FWHMs. From *top to bottom* and *left to right*, these show the effects of smoothing with kernels of the following FWHMs: 0, 2, 4, 6, 8, 10, 12, 14 and 16 mm

smoothing increases the signal-to-noise ratio in the data. The matched filter theorem stipulates that the optimal smoothing kernel corresponds to the size of the effect that one anticipates. A kernel similar in size to the anatomical extent of the expected haemodynamic response should therefore be chosen. Secondly, thanks to the central limit theorem, smoothing the data will render errors, or noise, more normally distributed, and will validate the use of inference based on parametric statistics. Thirdly, as we shall see later, when using the random field theory to make inference about regionally specific effects, there are specific assumptions that require smoothness in the data to be substantially greater than the voxel size (typically, as a rule of thumb, about three times the voxel size). Fourthly, small misregistration errors are inevitable in group studies; smoothing increases the degree of anatomical and functional overlap across subjects, reduces the effects of misregistration and thereby increases the significance of ensuing statistical tests.

In practice, there is no definitive amount of smoothing that should be applied to any data set; choice of smoothing kernel depends on the resolution of the data, the regions under investigation and single subject versus group analysis amongst other things. Commonly used FWHMs are between 6- and 12-mm isotropic.

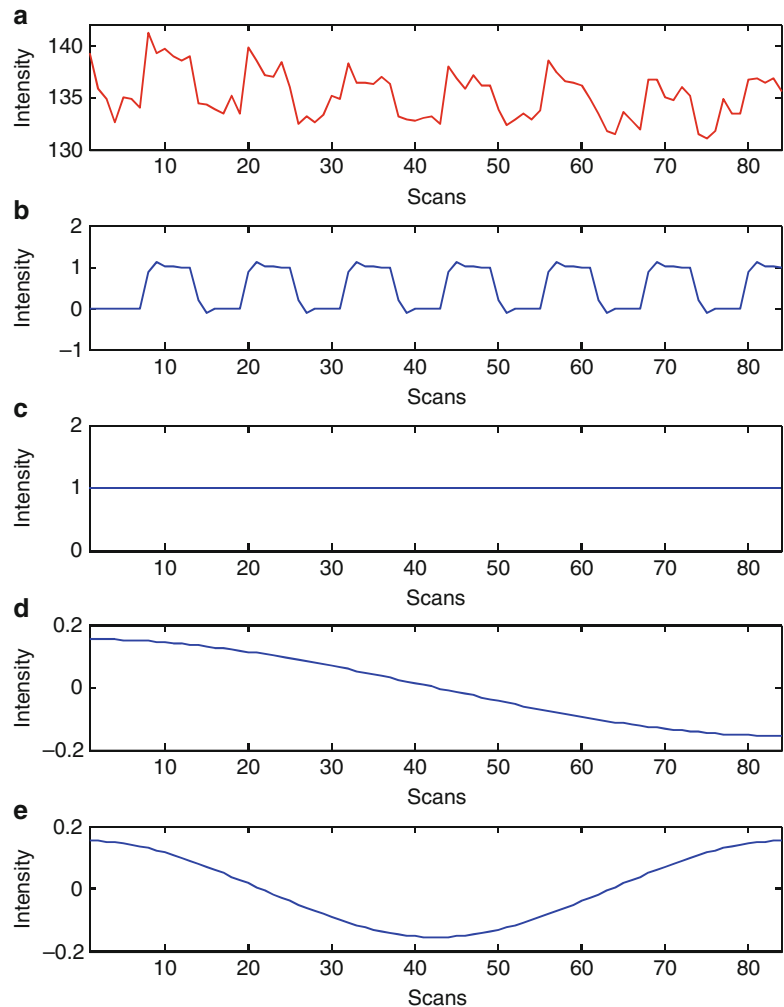
For the auditory data set, we will smooth the 84 normalised scans with a [6, 6, 6]-mm FWHM kernel to produce a new set of 84 images with an ‘s’ prefix.

6.4 Modelling and Statistical Inference

Statistical parametric mapping is a voxel by voxel hypothesis testing approach through which regions that show a significant experimental effect of interest are identified (Friston et al. 1991). It relies upon the construction of statistical parametric maps (SPMs), which are images with values at each voxel that are, under the null hypothesis, distributed according to a known probability density function (usually the Student’s t - or F -distributions). The parameters used to compute a standard univariate statistical test at each and every voxel in the brain are obtained from the estimation of a *general linear model* which partitions observed responses into components of interest (such as the experimental effect of interest), confounding factors (examples of such will be given later) and error (or ‘noise’) (Friston et al. 1994a). Hence, SPM is a *mass-univariate* approach: statistics are calculated independently at each voxel. The *random field theory* is then used to characterise the SPM and resolve the multiple comparisons problem induced by making inferences over a volume of the brain containing multiple voxels (Worsley et al. 1992, 1996; Friston et al. 1994b). ‘Unlikely’ topological features of the SPM, like activation peaks, are interpreted as regionally specific effects attributable to the experimental manipulation.

In this section, we will describe the general linear model in the context of fMRI time series. We will then estimate its parameters using the

Fig. 6.9 Predictors of an fMRI time series: **(a)** raw time series at a given voxel in the brain, **(b)** stimulus function convolved with the canonical haemodynamic response function, **(c)** constant term modelling the mean whole brain activity, **(d, e)** the first two components of a discrete cosine basis set modelling slow fluctuations



maximum likelihood method and describe how to test hypotheses by making statistical inferences on some of the parameter estimates by using *contrast*. The resulting statistical parameters are assembled into an image: this is the SPM. The random field theory provides adjusted p values to control false-positive rate for the search volume.

6.4.1 The General Linear Model

After the spatial preprocessing of the fMRI data, we can assume that all data from one particular voxel are derived from the same part of the brain, and that in any single subject, the data from that voxel form a sequential time series. A time series

selected from a (carefully chosen) voxel in the auditory data set is shown in Fig. 6.9a: variation in the response over time can be seen. There are 84 values, or data points, or observations. The aim is now to define a generative model of these data. This involves defining a prediction of what we might expect to observe in the measured BOLD signal given our knowledge of the acquisition apparatus and the experimental design. Here, the paradigm consisted of alternating periods (or ‘blocks’) of rest and auditory stimulation, with each block lasting for six scans. We expect that a voxel in a brain region sensitive to auditory stimuli will show a response that alternates with the same pattern and would thus, in the absence of noise, look like the plot in Fig. 6.10a.

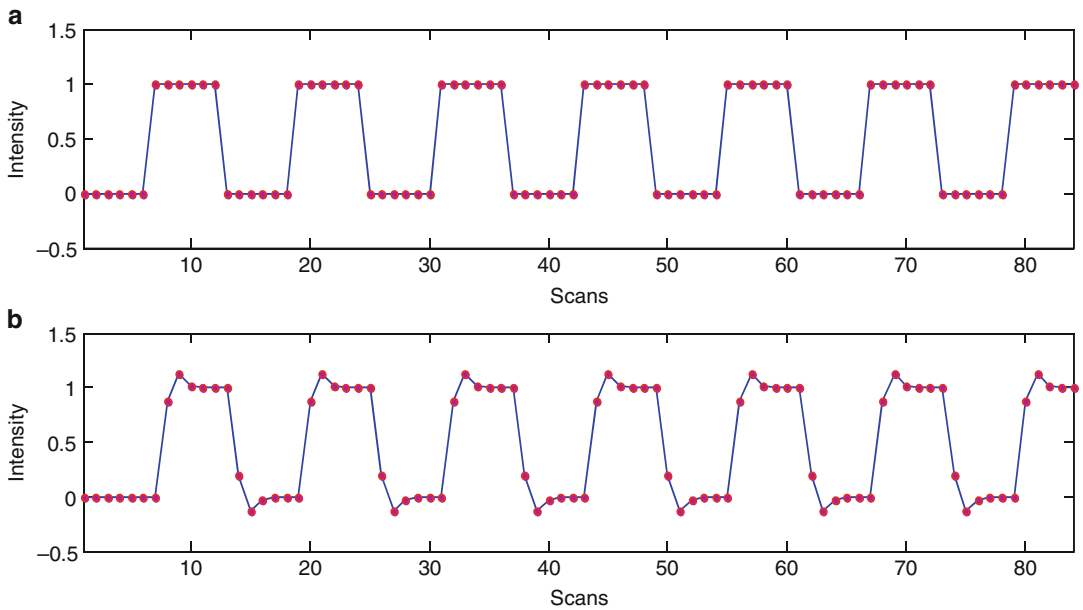
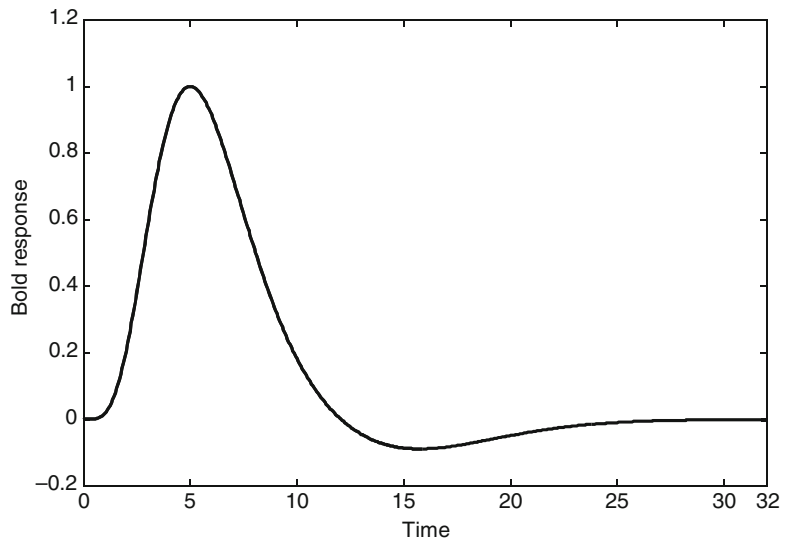


Fig. 6.10 Effect of convolution by the haemodynamic response function: (a) stimulus function constructed as a boxcar function and (b) after convolution with the canonical HRF

Fig. 6.11 Canonical haemodynamic response function (HRF) as used by SPM. This is the typical BOLD response to a single, impulse stimulation



However, we also know that with fMRI we are not directly measuring the neuronal activity, but the brain oxygen level-dependent (BOLD) signal with which it is associated. The observed BOLD signal corresponds to neuronally mediated haemodynamic change which can be modelled as a convolution of the underlying neuronal process by a *haemodynamic response function* (HRF). This function is called the *impulse response func-*

tion: it is the response that would be observed in the BOLD signal in the presence of a brief neuronal stimulation at $t=0$. The *canonical HRF* used in SPM is depicted in Fig. 6.11. The HRF models the fact that the BOLD response peaks about 5 s after the neuronal stimulation and takes about 32 s to go back to baseline, in a slow and smooth fashion, undershooting towards the end before reaching baseline.

We can thus improve our prediction by modifying the box car stimulus function of Fig. 6.10a to take into account the shape of the HRF. This is done through convolution, by assuming a linear time-invariant model. This convolution operation is conceptually the same as the one that was used in the smoothing preprocessing step; that was a convolution in space with a Gaussian kernel, whilst here it is a convolution in time with the canonical HRF. The output of this mathematical operation is displayed in Fig. 6.10b.

Looking at the raw time series of Fig. 6.9a, we can also directly observe that the mean of the signal is not zero; this should also be part of our prediction model. We model the (non-zero) mean of the signal with a predictor that is held constant over time, as shown in Fig. 6.9c.

Furthermore, we can also observe some slow fluctuations in the measured signal: what seems to be the response to the first block of stimulation has a higher amplitude than the response to the last one shown. There are indeed some low-frequency components in fMRI signals; these can be attributed to scanner drift (small changes in the magnetic field of the scanner over time) and/or to the effect of cardiac and respiratory cycles. As slow fluctuations are something that we expect in the data, we should also define predictors for them. A solution is to model the fluctuations through a discrete cosine transform basis set: a linear combination of cosine waves at several frequencies can accommodate a range of fluctuations. In order to remove any function with a cycle longer than 128 s (the default in SPM) and given the sampling rate and the number of scans, nine components are here required in the basis set. The first two components are displayed in Fig. 6.9d, e. Together, the set of cosine waves will effectively act as a high-pass filter with a 128-s cutoff.

Our best prediction of the observed data in Fig. 6.9a will then be a linear superposition of all the effects and confounds defined above and displayed in Fig. 6.9b–e. This is the assumption underlying the general linear model (GLM): the observed response (BOLD signal) y is expressed in terms of a linear combination of explanatory

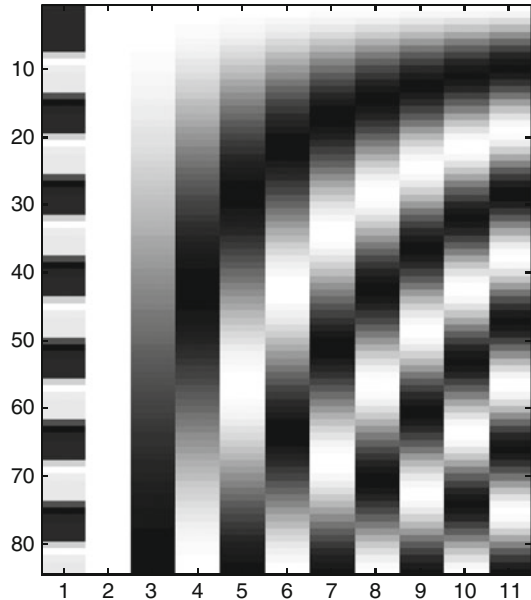


Fig. 6.12 Design matrix for the auditory data set: the first column models the condition-specific effect (boxcar function convolved with the HRF); the next column is a constant term, while the last nine columns are the components of a discrete cosine basis set modelling signal drifts over time. Note that a design matrix as displayed in SPM will not show the last nine terms

variables plus a well-behaved error term ε (Friston et al. 1994a):

$$y = X\beta + \varepsilon$$

The matrix X contains column-wise all the predictors that we have defined: everything we know about the experimental design and all potential confounds. This matrix is referred to as the *design matrix*. The one described so far is depicted in Fig. 6.12: it has 84 rows and 11 columns, each representing a predictor (or explanatory variable, covariate, regressor). This is just another way of representing conjointly the time series of Fig. 6.9 as an image where white represents a high value and black a low one.

The relative contribution of each of these columns to the response is controlled by the parameters β . These are the weights or regression coefficients of the GLM and will correspond to the size of the effects that we are measuring. β is

a vector whose length is the number of regressors in the design matrix, that is, its number of columns. The β -parameters are the unknown factor in this model.

Finally, the error term ε contains everything that cannot be explained by the model; these values are also known as the residuals, that is, the difference between the data y and the model prediction $X\beta$. In the simplest case, ε is assumed to follow a Gaussian distribution with a mean of zero and a standard deviation σ .

The general linear model is a very generic framework that encompasses many standard statistical analysis approaches: multiple regression, analysis of variance (ANOVA), analysis of covariance (ANCOVA) and t tests can all be framed in the context of a GLM and correspond to a particular form of the design matrix.

Fitting the GLM, or inverting the generative model, is the process of estimating its parameters given the data that we observed. This corresponds to adjusting the β -parameters of the model in order to obtain the best fit of the model to the data. Another way of thinking of this is that we need to find the β -parameters that minimise the error term ε . It can be shown that under the assumption that the errors are normally distributed, the parameters can be estimated using the following equation:

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

This is the ordinary least squares (OLS) equation that relates the estimated parameters $\hat{\beta}$ to the design matrix X and the observed time series y .

Figure 6.13 shows how the GLM with the design matrix shown in Fig. 6.12 fits the time series shown in Fig. 6.9a and reproduced in Fig. 6.13a in blue. The predicted time series is overlaid in red; it is a linear combination of the stimulus function (Fig. 6.13b), the mean whole brain activity (Fig. 6.13c) and the low-frequency drifts (Fig. 6.13d). The residuals are displayed in Fig. 6.13e; they are the difference between the observed time series and its model prediction.

This procedure is repeated for all voxels within the brain, generating maps of the estimated regression coefficients $\hat{\beta}$. The variance of the noise $\hat{\sigma}^2$ is also estimated voxel-wise. As mentioned above, this is essentially a mass-univariate approach: the same model (design matrix X) is fitted independently to the time series at every voxel, providing local estimates of the effect sizes.

Following on from this description of how data are modelled, there are a few more considerations that need to be taken into account:

- In practice, the low-frequency components from the discrete cosine transform (DCT) basis set are not added to the design matrix, but the data and design are instead temporally filtered before the model is estimated. This is mathematically identical but computationally more efficient as drift effects will always be confounds of no interest that are not tested. Hence, in practice they do not appear in the design matrix as in Fig. 6.12, but they are still dealt within the background (and the degrees of freedom of the model are adjusted accordingly).
- When using the canonical HRF to model the transfer function from neuronal activity to BOLD response, we assumed that it was known and fixed. However, in practice this is not actually the case; the HRF varies across brain regions and across individuals. A solution is to use a set of basis functions rather than a single function in order to add some flexibility to the modelling of the response. The HRF will then be modelled as a linear combination of these basis functions. A popular choice, providing flexibility and parsimony, is to use the *informed* basis set: this consists of the canonical HRF and its temporal and dispersion derivatives, as shown in Fig. 6.14. According to the weight given to each of the components, the informed basis set allows us to model a shift in the latency of the response (with the temporal derivative) and changes in the width of the response (with the dispersion derivative). When using the informed basis set, each experimental condition is modelled

Fig. 6.13 Fit of the GLM defined earlier on the fMRI time series of Fig. 6.9: (a) observed fMRI time series in *blue* and model prediction in *red*, fitted predictors for the (b) condition-specific effect, (c) constant term, (d) slow frequency fluctuations and (e) residuals

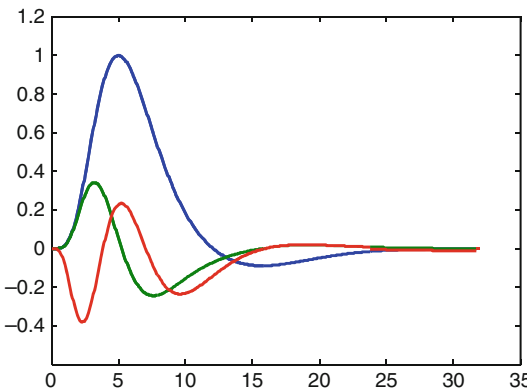
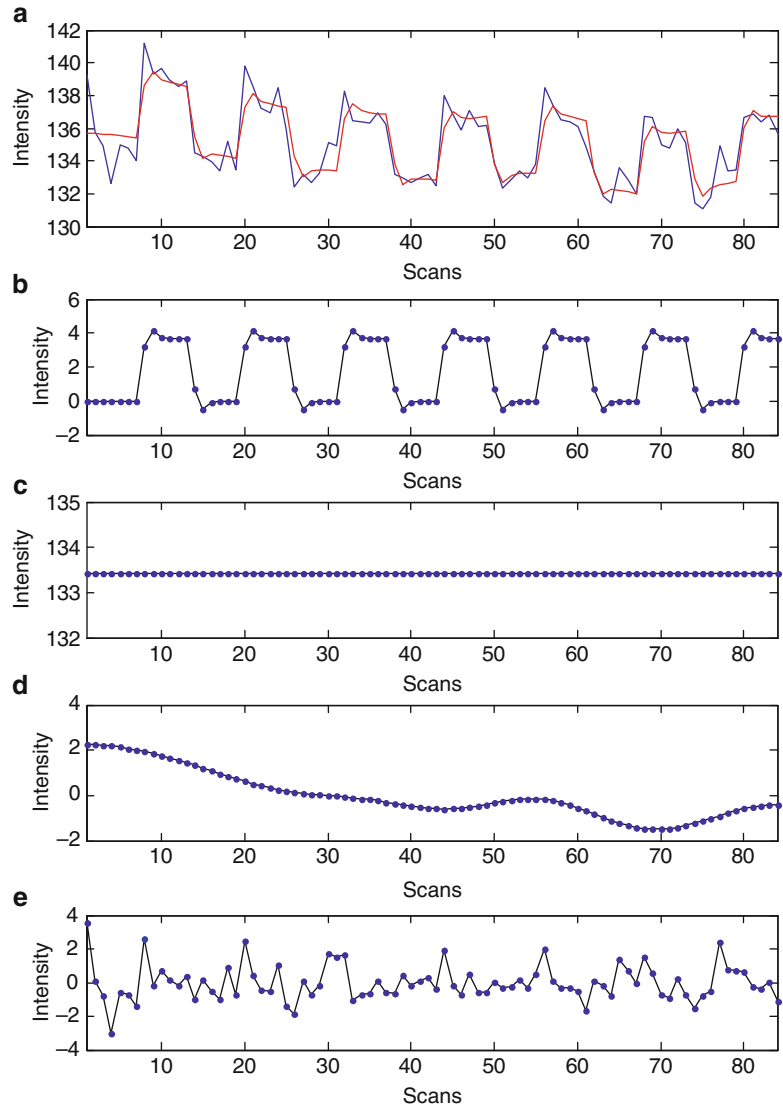


Fig. 6.14 The ‘informed’ basis set: the canonical HRF (*blue*) and its temporal (*green*) and dispersion (*red*) derivatives

by a set of three regressors, each of which is the neuronal activity stimulus function convolved separately with one of the three components. The predicted response for that condition will be a linear condition of these three regressors. The temporal derivative is also useful to model slice-timing issues. In multislice acquisitions, different slices are acquired at different times. A solution is to temporally realign the data as if they were acquired at the same time through interpolation. This is called slice-timing correction and is a possible option during the preprocessing of the data. Using the informed basis set is an

alternative way to correct for the same effect (please see (Sladky et al. 2011) for a recent comparison of the two approaches).

- fMRI data exhibit short-range serial or temporal correlations. This means that the error at time t is correlated with the error at previous time points. This has to be modelled, because ignoring these correlations may lead to invalid statistical testing. An error covariance matrix must therefore be estimated by assuming some kind of non-sphericity, a departure from the independent and identically distributed assumptions of the noise (Worsley and Friston 1995). A popular model used to capture the typical form of serial correlation in fMRI data is the autoregressive model of order 1, AR(1), relating the error at time t to the error at time $t-1$ with a single parameter. It can be estimated efficiently and precisely by pooling its estimate over voxels. Once the error covariance matrix is estimated, the GLM can be inverted using weighted least squares (WLS) instead of OLS; alternatively, the estimated error covariance matrix can be used to whiten data and design, that is, to *undo* the serial correlation, so that OLS can be applied again. This is the approach implemented in SPM.
- It is also possible to add regressors to the design matrix without going through the convolution process described above. An important example is the modelling of residual movement-related effects. Because movement expresses itself in the data directly and not through any haemodynamic convolution, it can be added directly as a set of explanatory variables. Similarly, in the presence of an abrupt artefact in the data corrupting one scan, a strategy is to model it as a regressor that is zero everywhere but one at that scan. This will effectively covary out that artefactual value in the time series, reducing the inflated variance that it was contributing to. This is better than manually removing that scan prior to analysis as it preserves the temporal process.
- An important distinction in experimental design for fMRI is that between event- and epoch-related designs. Event-related fMRI is simply the use of fMRI to detect responses to individual trials (Josephs et al. 1997). The neuronal activ-

ity is usually modelled as a delta function – an *event* – at the trial onset. Practically speaking, in SPM we assume that the duration of a trial is zero. In an epoch-related design, however, we assume that the duration of the trial is greater than zero. This is the case in block-design studies, in which the responses to a sequence of trials (which all evoke the same experimental effect of interest) are modelled together (as an epoch). There are otherwise no conceptual changes in the statistical analysis of event-related and epoch-related (block) designs. One of the advantages of event-related designs is that trials of different types can be intermixed instead of blocking events of the same type together, allowing the measurement of a greater range of psychological effects. There are a number of considerations which impact on the choice of an experimental design, including the constraints imposed by high-pass filtering and haemodynamic convolution of the data affecting its efficiency. We refer interested readers to Chapter 15 of (Friston et al. 2007) or its online version¹¹ for a thorough examination of design efficiency.

For the auditory data set, the first step is to specify the design matrix; this is done through the ‘Specify first-level’ button. After specifying a directory in which the results will be stored, the inputs to specify are the units in which the onsets and duration of each trial will be entered (these can be either ‘scans’ or ‘seconds’; we will use ‘scans’ in this example), the TR (7 s) and the actual preprocessed data to be analysed (the 84 files with an ‘sw’ prefix). In this data set, there is just one condition to specify: the onsets are [6 18 30 42 54 66 78], corresponding to the scan number at the beginning of each auditory stimulation block, and the durations are [6], indicating that each auditory stimulation block lasts for six scans (with a rest block in between each one). The movement parameters can be added as extra regressors using the ‘Multiple regressors’ entry by selecting the ‘rp_*.txt’ file that was saved during the realignment. Other parameters can be left as default, especially the high-pass filter cut-off

¹¹<http://imaging.mrc-cbu.cam.ac.uk/imaging/DesignEfficiency>

(128 s), the use of the canonical HRF only and the modelling of serial correlation using an AR(1) model. The output is an *SPM.mat* file; this contains all the information about the data and the model design. The design matrix is also displayed for review. As expected, it has eight columns: the first column is the block stimulus function convolved by the HRF, the following six columns are the movement parameters (three translations and three rotations, see Fig. 6.5) and the last column is a constant term modelling the whole brain activity. The ‘Estimate’ button then allows us to invert this GLM and estimate its parameters. A number of image files will be created, including eight maps of the estimated regression coefficients, one for each column of the design matrix (*beta_*.{hdr,img}*) and one mask image (*mask.{hdr,img}*), which contains a binary volume indicating which voxels were included in the analysis.

6.4.2 Contrasts

Having specified and estimated parameters of the general linear model, the next step is to make a statistical inference about those parameters. This is done by using their estimated variance. Some of the parameters will be of interest (those pertaining to the experimental conditions), while others will be of no interest (those pertaining to confounding effects). Inference allows one to test the null hypothesis that all the estimates are zero, using the *F*-statistic to give an *SPM{F}*, or that some particular linear combination (e.g. a subtraction) of the estimates is zero, using the *t*-statistic to give an *SPM{t}*. A linear combination of regression coefficients is called a *contrast*, and its corresponding vector of weights *c* is called a contrast vector.

The *t*-statistic is obtained by dividing a contrast (specified by contrast weights) of the associated parameter estimates by the standard error of that contrast. The latter is estimated using the variance of the residuals $\hat{\sigma}^2$.

$$T = \frac{c^T \hat{\beta}}{\sqrt{\hat{\sigma}^2 c^T (X^T X)^{-1} c}}$$

This is essentially a signal-to-noise ratio, comparing an effect size with its precision.

An example of a contrast vector would be $c^T = [1 \ -1 \ 0 \dots]$ to compare the difference in responses evoked by two conditions, modelled by the first two condition-specific regressors in the design matrix. In SPM, a *t* test is signed, in the sense that a contrast vector $c^T = [1 \ -1 \ 0 \dots]$ is looking for a greater response in the first condition than in the second condition, while a contrast $c^T = [-1 \ 1 \ 0 \dots]$ is looking for the opposite effect. In other words, it means that a *t*-contrast tests the null hypothesis $c^T \beta = 0$ against the one-sided alternative $c^T \beta > 0$. The resulting *SPM{t}* is a statistic image, with each voxel value being the value of the *t*-statistic for the specified contrast at that location. Areas of the *SPM{t}* with high voxel values (higher than one might expect by chance) indicate evidence for ‘neural activations’.

Similarly, if you have a design where the third column in the design matrix is a covariate, then the corresponding parameter is essentially a regression slope, and a contrast with weights $c^T = [0 \ 0 \ 1 \ 0 \dots]$ tests the hypothesis of zero regression slope, against the alternative of a positive slope. This is equivalent to a test of no correlation, against the alternative of positive correlation. If there are other terms in the model beyond a constant term and the covariate, then this correlation is a partial correlation, the correlation between the data and the covariate after accounting for the other effects (Andrade et al. 1999).

Sometimes, several contrasts of parameter estimates are jointly interesting. In these instances, an *SPM{F}* is used and is specified with a matrix of contrast weights which can be thought of as a collection of *t*-contrasts. An *F*-contrast might look like this

$$c^T = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots \\ 0 & -1 & 0 & 0 & \dots \end{bmatrix}$$

This would test for the significance of the first or the second parameter estimates. The fact that the second weight is negative has no effect on the test because the *F*-statistic is blind to sign as it is based on sums of squares. The *F*-statistic can

also be interpreted as a model comparison device, comparing two nested models using the extra sum-of-squares principle. For the F -contrast above, this corresponds to comparing the specified full model with a reduced model where the first two columns would have been removed. F -contrasts are mainly used either as two-sided tests (the $SPM\{F\}$ then being the square of the corresponding $SPM\{t\}$) or to test the significance of effects modelled by several columns. Effects modelled by several columns might include the use of a multiple basis set to model the HRF, a polynomial expansion of a parametric modulated response or a contrast testing more than two levels in a factorial design.

As with the $SPM\{t\}$, the resulting $SPM\{F\}$ is a statistic image, with voxel values the value of the F -statistic for the specified contrast at that location. Areas of the $SPM\{F\}$ with high voxel values indicate evidence for ‘neural activations’.

6.4.3 Topological Inference

Having computed the statistic, we need to decide whether it represents convincing evidence of the effect in which we are interested; this decision is the process of making a statistical inference. This is done by testing the statistic against the null hypothesis that there is no effect. Here, the null hypothesis is distributed according to a known parametric probability density function, a Student’s t - or F -distribution. Then, by choosing a significance level (which is the level of control over the false-positive error rate, usually chosen as 0.05), we can derive a critical threshold above which we will reject the null hypothesis and accept the alternative hypothesis that there is convincing evidence of an effect. If the observed statistic is lower than the critical threshold, we fail to reject the null hypothesis and we must conclude that there is no convincing evidence of an effect. A p value can also be computed to measure the evidence against the null hypothesis: this is the probability of observing a statistic at least as large as the one observed under the null hypothesis (i.e. by chance).

The problem we face in functional imaging is that we are not dealing with a single statistic value, but with an image that comprises many thousands of voxels and their associated statistical values. This gives rise to the multiple comparisons problem, which is a consequence of the use of a mass-univariate approach: as a general rule, without using an appropriate method of correction, the greater the number of voxels tested, the greater the number of false positives. This is clearly unacceptable and requires the definition of a new null hypothesis which takes into account the whole volume, or family, of statistics contained in an image: the *family-wise* null hypothesis that there is no effect *anywhere* in the entire search volume (e.g. the brain). We then aim to control the family-wise error rate (FWER) – the probability of making one or more false positives over the entire search volume. This results in adjusted p values, *corrected* for the search volume.

A traditional statistical method for controlling FWER is to use the Bonferroni correction, in which a voxel-wise significance level simply corresponds to the family-wise significance level (e.g. 0.05) divided by the number of tests (i.e. voxels). However, this approach assumes that every test (voxel statistic) is independent and is too conservative to use in the presence of correlation between tests, such as the case with functional imaging data. Functional imaging data are intrinsically smoothed due to the acquisition process and have also been smoothed as part of the spatial preprocessing; neighbouring voxel statistics are therefore not independent. The random field theory (RFT) provides a way of adjusting the p value to take this into account (Worsley et al. 1992; Friston et al. 1994b; Worsley et al. 1996). Providing that data are smooth, the RFT adjustment is less severe (i.e. more sensitive) than a Bonferroni correction for the number of voxels. The p value is a function of the search volume and *smoothness* (parameterised as the FWHM of the Gaussian kernel required to simulate images with the same apparent spatial smoothness as the one we observe). A description of the random field theory is well beyond the scope of this chapter, but it is worth mentioning that one of the

assumptions for its application on discrete data fields is that the observed fields are smooth. This was one of the motivations for smoothing the fMRI data as a preprocessing step. In practice, smoothness will be estimated from the data themselves (to take into account both intrinsic and explicit smoothness) (Kiebel et al. 1999). The RFT correction discounts voxel sizes by expressing the search volume in terms of smoothness or resolution elements (*resels*).

To make inferences about regionally specific effects, the SPM is thresholded using height and spatial extent thresholds that are specified by the user. Corrected p values can then be derived that pertain to topological features of the thresholded map (Friston et al. 1996a):

- The number of activated regions (i.e. the number of clusters above the height and volume threshold). These are *set-level* inferences.
- The number of activated voxels (i.e. the volume or extent) comprising a particular cluster. These are *cluster-level* inferences.
- The height of each local maxima, or peak, within that cluster. These are *peak-level* inferences.

Set-level inferences are generally more powerful than cluster-level inferences, which are themselves generally more powerful than peak-level inferences. The price paid for this increased sensitivity is a reduced localising power. Peak-level tests permit individual maxima to be identified as significant, whereas cluster and set-level inferences only allow clusters or a set of clusters to be declared significant. In some cases, however, focal activation might actually be detected with greater sensitivity using tests based on peak height (with a spatial extent threshold of zero). In practice, this is the most commonly used level of inference, reflecting the fact that characterisation of functional anatomy is generally more useful when specified with a high degree of anatomical precision.

When making inferences about regional effects in SPMs, one often has some idea *a priori* about where the activation should be. In this instance, a correction for the entire search volume is inappropriately stringent. Instead, a small search volume within which analyses will be carried out can be specified beforehand, and an RFT

correction applied restricted to that region only (Worsley et al. 1996). This is often referred to as a small volume correction (SVC).

For the auditory data set, inference is performed through the ‘Results’ button by selecting the *SPM.mat* file from the previous step. To test for the positive effect of passive listening to words versus rest, the t -contrast to enter is $c^T = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$. Two files will be created on disk at this stage: a *con_0001.{hdr,img}* file of the contrast image (here identical to *beta_0001.{hdr,img}*) and the corresponding *SPM{T} spmT_0001.{hdr,img}*. Choosing a 0.05 FWE-corrected threshold yields the results displayed in Fig. 6.15. The maximum intensity projection (MIP) image gives an overview of the activated regions, the auditory cortices. This can be overlaid on the anatomy of the subject (the normalised ‘*wm*.img*’ image file) using the menu entry ‘Overlays>Sections’ of the Interactive window. The button ‘whole brain’ will display the results table (see Fig. 6.16) listing the p values adjusted for the search volume (the whole brain here) for all topological features of the excursion set: local maxima height for peak-level inference, cluster extent for cluster-level inference and number of clusters for set-level inference. The footnote of the results table lists some numbers pertaining to the RFT: the estimated FWHM smoothness is [10.9, 10.9, 9.2] in mm and the number of resels is 1,450 here. These are the values reflecting the size and smoothness of the search volume that are used to control for the FWER.

Guidelines for reporting an fMRI study in a publication are given in Poldrack et al. (2008).

6.4.4 Population-Level Inference

Neuroimaging data from multiple subjects can be analysed using fixed-effects (FFX) or random-effects (RFX) analyses (Holmes and Friston 1998). FFX analysis is used for reporting case studies, while RFX is used to make inferences about the population from which the subjects were drawn. In the former, the error variance is estimated on a scan-by-scan basis and contains contributions from within-subject variance only. We are therefore not

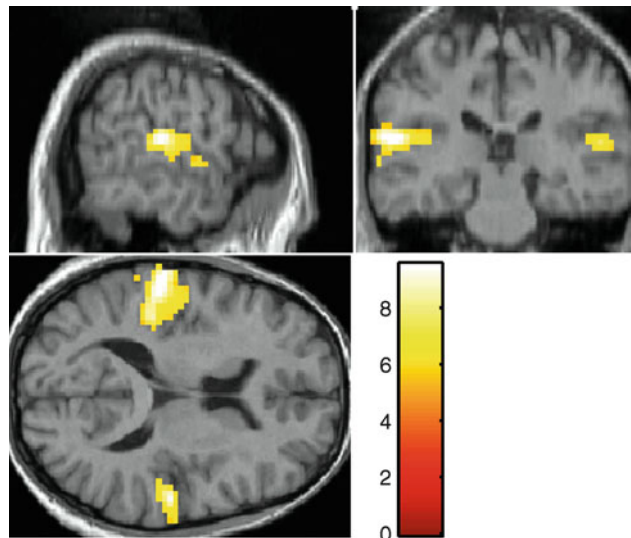
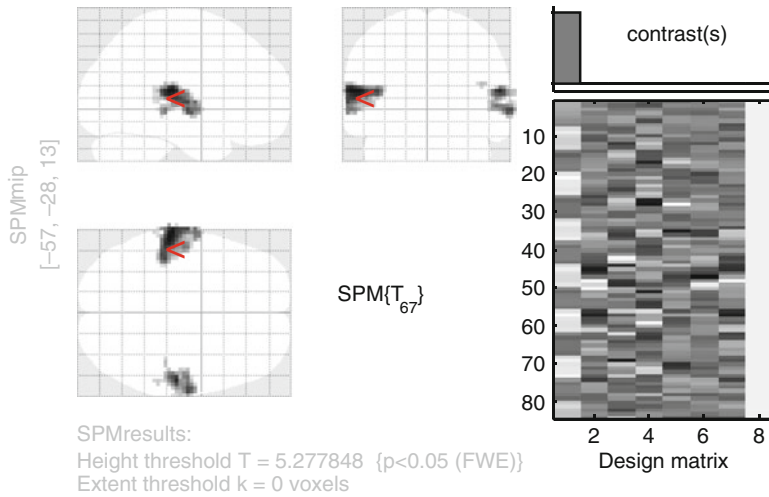


Fig. 6.15 Results of the statistical inference for the auditory data set when looking for regions showing an increased activity when words are passively listened to in comparison with rest. *Top right:* design matrix and contrast $c^T = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$. *Top left:*

$SPM\{t\}$ displayed as a maximum intensity projection over three orthogonal planes. To control for $p < 0.05$ corrected, the applied threshold was $T = 5.28$. *Lower panel:* Thresholded $SPM\{t\}$ overlaid on the normalised structural scan of that subject, highlighting the bilateral activation

making formal inference about population effects using FFX, but are restricted to informal inferences based on separate case studies or summary images showing the average group effect. (This is implemented in SPM by concatenating data from all subjects into the same GLM, simply modelling subjects as coming from separate sessions.) Conversely, random-effects analyses take into

account both sources of variance (within- and between-subject). The term ‘*random effect*’ indicates that we have accommodated the randomness of different responses from subject to subject.

Both analyses are perfectly valid but only in relation to the inferences that are being made: inferences based on fixed-effect analyses are about the particular subjects studied. Random-effects

Statistics: p -values adjusted for search volume

set-level		cluster-level				peak-level					mm mm mm		
p	c	$p_{\text{FWE-corr}}$	$q_{\text{FDR-corr}}$	k_E	p_{uncorr}	$p_{\text{FWE-corr}}$	$q_{\text{FDR-corr}}$	T	$(Z_{\text{=}})$	p_{uncorr}			
0.000	5	0.000	0.000	328	0.000	0.000	0.000	9.58	7.58	0.000	-57	-28	13
						0.000	0.000	8.34	6.88	0.000	-42	-31	13
						0.000	0.000	8.31	6.87	0.000	-66	-13	4
		0.000	0.000	151	0.000	0.000	0.000	8.97	7.25	0.000	57	-22	13
						0.000	0.000	8.26	6.84	0.000	66	-13	-2
						0.000	0.001	7.38	6.29	0.000	54	-13	1
		0.002	0.072	4	0.043	0.003	0.106	6.02	5.36	0.000	42	-31	19
		0.015	0.288	1	0.288	0.034	0.757	5.38	4.89	0.000	69	-25	-2
		0.015	0.288	1	0.288	0.039	0.800	5.35	4.86	0.000	72	-34	-8

table shows 3 local maxima more than 8.0mm apart

Height threshold: $T = 5.28$, $p = 0.000$ (0.050)
 Extent threshold: $k = 0$ voxels, $p = 1.000$ (0.050)
 Expected voxels per cluster, $\langle k \rangle = 0.957$
 Expected number of clusters, $\langle c \rangle = 0.05$

Degrees of freedom = [1.0, 67.0]
 FWHM = 10.9 10.9 9.2 mm mm mm; 3.6 3.6 3.1 {voxels}
 Volume: 1778922 = 65886 voxels = 1449.9 resels
 Voxel size: 3.0 3.0 3.0 mm mm mm; (resel = 40.60 voxels)

Fig. 6.16 Results table summarising the inferences made with the t -contrast $c^T = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$, adjusted for the search volume (the whole brain in this case). See text for description of the key elements

analyses are usually more conservative but allow the inference to be generalised to the population from which the subjects were drawn.

In practice, RFX analyses can be implemented using the computationally efficient ‘*summary-statistic*’ approach. Contrasts of parameters estimated from a *first-level* (within-subject) analysis are entered into a *second-level* (between-subject) analysis. The second-level design matrix then simply tests the null hypothesis that the contrasts are zero (and is usually a column of ones, implementing a one-sample t test). The validity of the approach rests upon the use of balanced designs (all subjects have identical design matrices and error variances) but has been shown to be remarkably robust to violations of this assumption (Mumford and Nichols 2009).

For our auditory data set, if we had scanned 12 subjects, for example, each of whom performed the same task, the group analysis would entail (i) applying the same spatial preprocessings to each of the 12 subjects, (ii) fitting a first-level GLM independently to each of the 12 subjects, (iii) defining the effect of interest for each subject with a contrast vector $c^T = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$ and producing a contrast image containing the contrast of the parameter estimates at each

voxel and (iv) feeding each of the 12 contrast images into a second-level GLM, through which a one-sample t test could be carried out across all 12 subjects to find the activations that show significant evidence of a population effect.

6.5 Conclusions

In this chapter we have described how statistical parametric mapping can be used to identify and characterise regionally specific effects in functional MRI data. We have also illustrated the principles of SPM through the analysis of a block-design data set using the SPM software. After preprocessing the data to correct them for movement and normalise them into a standard space, the general linear model and random field theory are used to analyse and make classical inferences. The GLM is used to model BOLD responses to given experimental manipulations. The estimated parameters of the GLM are used to compute a standard univariate statistical test at each and every voxel, leading to the construction of statistical parametric maps. The random field theory is then used to resolve the multiple comparisons problem induced by inferences over a

volume of the brain containing many voxels. RFT provides a method for adjusting p values for the search volume of a statistical parametric map to control false-positive rates.

We have here described the fundamental methods used to carry out fMRI analyses in SPM. There are, however, many additional approaches and tools that can be used to refine and extend analyses, such as voxel-based morphometry (VBM) to analyse structural data sets (Ashburner and Friston 2000; Ridgway et al. 2008) and dynamic causal modelling (DCM) to study effective connectivity (Friston et al. 2003; Seghier et al. 2010; Stephan et al. 2010). While the key steps of the SPM approach we describe above remain broadly constant, SPM software (along with many other software analysis packages) is constantly evolving to incorporate advances in neuroimaging analysis; we encourage readers to explore these exciting new developments.

References

- Andersen SM, Rapcsak SZ et al (2010) Cost function masking during normalization of brains with focal lesions: still a necessity? *Neuroimage* 53:78–84. doi:10.1016/j.neuroimage.2010.06.003
- Andersson JLR, Hutton C et al (2001) Modeling geometric deformations in EPI time series. *Neuroimage* 13:903–919. doi:10.1006/nimg.2001.0746
- Andrade A, Paradis AL et al (1999) Ambiguous results in functional neuroimaging data analysis due to covariate correlation. *Neuroimage* 10:483–486. doi:10.1006/nimg.1999.0479
- Ashburner J (2011) SPM: a history. *Neuroimage*. doi:10.1016/j.neuroimage.2011.10.025
- Ashburner J, Friston KJ (1999) Nonlinear spatial normalization using basis functions. *Hum Brain Mapp* 7:254–266
- Ashburner J, Friston KJ (2000) Voxel-based morphometry – the methods. *Neuroimage* 11:805–821. doi:10.1006/nimg.2000.0582
- Ashburner J, Friston KJ (2005) Unified segmentation. *Neuroimage* 26:839–851. doi:10.1016/j.neuroimage.2005.02.018
- Brett M, Leff AP et al (2001) Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage* 14:486–500. doi:10.1006/nimg.2001.0845
- Brett M, Johnsrude IS et al (2002) The problem of functional localization in the human brain. *Nat Rev Neurosci* 3:243–249. doi:10.1038/nrn756
- Collignon A, Maes F et al (1995) Automated multimodality image registration based on information theory. In: *Proceedings of information processing in medical imaging (IPMI)*, Ile de Berder, 1995
- Crinion J, Ashburner J et al (2007) Spatial normalization of lesioned brains: performance evaluation and impact on fMRI analyses. *Neuroimage* 37:866–875. doi:10.1016/j.neuroimage.2007.04.065
- Fox PT (1995) Spatial normalization origins: objectives, applications, and alternatives. *Hum Brain Mapp* 3:161–164. doi:10.1002/hbm.460030302
- Friston KJ, Frith CD et al (1991) Comparing functional (PET) images: the assessment of significant change. *J Cereb Blood Flow Metab* 11:690–699. doi:10.1038/jcbfm.1991.122
- Friston KJ, Holmes AP et al (1994a) Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 2:189–210. doi:10.1002/hbm.460020402
- Friston KJ, Worsley KJ et al (1994b) Assessing the significance of focal activations using their spatial extent. *Hum Brain Mapp* 1:210–220. doi:10.1002/hbm.460010306
- Friston KJ, Ashburner J et al (1995) Spatial registration and normalization of images. *Hum Brain Mapp* 3:165–189. doi:10.1002/hbm.460030303
- Friston KJ, Holmes A et al (1996a) Detecting activations in PET and fMRI: levels of inference and power. *Neuroimage* 4:223–235. doi:10.1006/nimg.1996.0074
- Friston KJ, Williams S et al (1996b) Movement-related effects in fMRI time-series. *Magn Reson Med* 35:346–355
- Friston KJ, Harrison L et al (2003) Dynamic causal modelling. *Neuroimage* 19:1273–1302
- Friston K, Ashburner J et al (2007) *Statistical parametric mapping: the analysis of functional brain images*. Elsevier/Academic, Amsterdam/Boston
- Holmes A, Friston K (1998) Generalisability, random effects and population inference. *Neuroimage* 7:S754
- Jezzard P, Balaban RS (1995) Correction for geometric distortion in echo planar images from B0 field variations. *Magn Reson Med* 34:65–73
- Josephs O, Turner R et al (1997) Event-related fMRI. *Hum Brain Mapp* 5:243–248
- Kiebel SJ, Poline JB et al (1999) Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. *Neuroimage* 10:756–766. doi:10.1006/nimg.1999.0508
- Mazziotta JC, Toga AW et al (1995) A probabilistic atlas of the human brain: theory and rationale for its development. The international consortium for brain mapping (ICBM). *Neuroimage* 2:89–101
- Mumford JA, Nichols T (2009) Simple group fMRI modeling and inference. *Neuroimage* 47:1469–1475. doi:10.1016/j.neuroimage.2009.05.034
- Poldrack RA, Fletcher PC et al (2008) Guidelines for reporting an fMRI study. *Neuroimage* 40:409–414. doi:10.1016/j.neuroimage.2007.11.048
- Ridgway GR, Henley SMD et al (2008) Ten simple rules for reporting voxel-based morphometry studies.

- Neuroimage 40:1429–1435. doi:[10.1016/j.neuroimage.2008.01.003](https://doi.org/10.1016/j.neuroimage.2008.01.003)
- Seghier ML, Zeidman P et al (2010) Identifying abnormal connectivity in patients using dynamic causal modeling of fMRI responses. *Front Syst Neurosci*. doi:[10.3389/fnsys.2010.00142](https://doi.org/10.3389/fnsys.2010.00142)
- Sladky R, Friston KJ et al (2011) Slice-timing effects and their correction in functional MRI. *Neuroimage* 58:588–594. doi:[10.1016/j.neuroimage.2011.06.078](https://doi.org/10.1016/j.neuroimage.2011.06.078)
- Stephan KE, Penny WD et al (2010) Ten simple rules for dynamic causal modeling. *Neuroimage* 49:3099–3109. doi:[10.1016/j.neuroimage.2009.11.015](https://doi.org/10.1016/j.neuroimage.2009.11.015)
- Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain: an approach to medical cerebral imaging. Thieme Medical, New York
- Unser M, Aldroubi A et al (1993) B-spline signal processing. I. Theory. *IEEE Trans Signal Process* 41:821–833. doi:[10.1109/78.193220](https://doi.org/10.1109/78.193220)
- Wells WM 3rd, Viola P et al (1996) Multi-modal volume registration by maximization of mutual information. *Med Image Anal* 1:35–51
- Worsley KJ, Friston KJ (1995) Analysis of fMRI time-series revisited – again. *Neuroimage* 2:173–181. doi:[10.1006/nimg.1995.1023](https://doi.org/10.1006/nimg.1995.1023)
- Worsley K, Evans A et al (1992) A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab* 12:900–918
- Worsley KJ, Marrett S et al (1996) A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Mapp* 4:58–73