1

Introduction

The goal of this book is to provide the reader with a solid background in the techniques used for processing and analysis of functional magnetic resonance imaging (fMRI) data.

1.1 A brief overview of fMRI

Since its development in the early 1990s, fMRI has taken the scientific world by storm. This growth is easy to see from the plot of the number of papers that mention the technique in the PubMed database of biomedical literature, shown in Figure 1.1. Back in 1996 it was possible to sit down and read the entirety of the fMRI literature in a week, whereas now it is barely feasible to read all of the fMRI papers that were published in the previous week! The reason for this explosion in interest is that fMRI provides an unprecedented ability to safely and noninvasively image brain activity with very good spatial resolution and relatively good temporal resolution compared to previous methods such as positron emission tomography (PET).

1.1.1 Blood flow and neuronal activity

The most common method of fMRI takes advantage of the fact that when neurons in the brain become active, the amount of blood flowing through that area is increased. This phenomenon has been known for more than 100 years, though the mechanisms that cause it remain only partly understood. What is particularly interesting is that the amount of blood that is sent to the area is more than is needed to replenish the oxygen that is used by the activity of the cells. Thus, the activity-related increase in blood flow caused by neuronal activity leads to a relative surplus in local blood oxygen. The signal measured in fMRI depends on this change in oxygenation and is referred to as the blood oxygenation level dependent, or BOLD, signal.

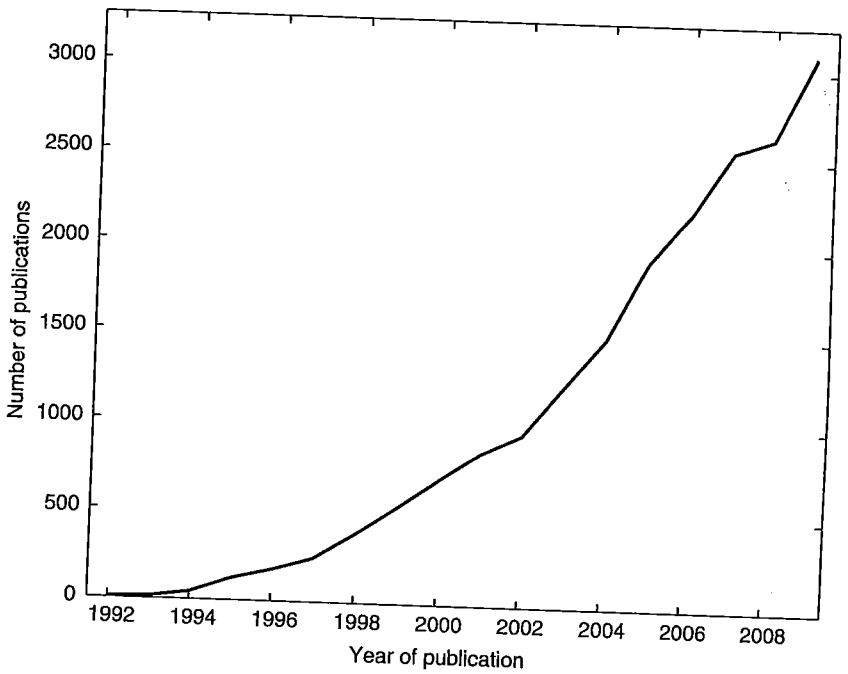

Figure 1.1. A plot of the number of citations in the PubMed database matching the query ["fMRI" O "functional MRI" OR "functional magnetic resonance imaging"] for every year since 1992.

Figure 1.2 shows an example of what is known as the *hemodynamic response*, whic is the increase in blood flow that follows a brief period of neuronal activity. Ther are two facts about the hemodynamic response that underlie the basic features o BOLD fMRI and determine how the data must be analyzed. First, the hemodynami response is slow; whereas neuronal activity may only last milliseconds, the increase i blood flow that follows this activity takes about 5 seconds to reach its maximum. Thi peak is followed by a long undershoot that does not fully return to baseline for at leas 15–20 seconds. Second, the hemodynamic response can, to a first approximation be treated as a *linear time-invariant* system (Cohen, 1997; Boynton et al., 1996 Dale, 1999). This topic will be discussed in much greater detail in Chapter 5, bu in essence the idea is that the response to a long train of neuronal activity can be determined by adding together shifted versions of the response to a shorter train of activity. This linearity makes it possible to create a straightforward statistical model that describes the timecourse of hemodynamic signals that would be expected given some particular timecourse of neuronal activity, using the mathematical operation of *convolution*.

1.1.2 Magnetic resonance imaging

The incredible capabilities of magnetic resonance imaging (MRI) can hardly be overstated. In less than 10 minutes, it is possible to obtain images of the human

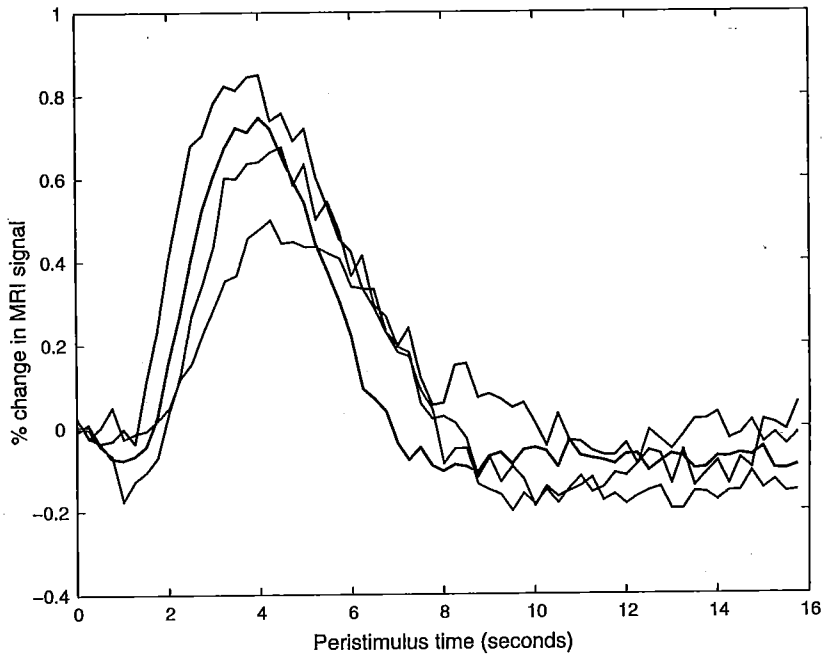

Figure 1.2. An example of the hemodynamic responses evoked in area V1 by a contrast-reversing checkerboard displayed for 500 ms. The four different lines are data from four different individuals, showing how variable these responses can be across people. The MRI signal was measured every 250 ms, which accounts for the noisiness of the plots. (Data courtesy of Stephen Engel, University of Minnesota)

brain that rival the quality of a postmortem examination, in a completely safe and noninvasive way. Before the development of MRI, imaging primarily relied upon the use of ionizing radiation (as used in X-rays, computed tomography, and positron emission tomography). In addition to the safety concerns about radiation, none of these techniques could provide the flexibility to image the broad range of tissue characteristics that can be measured with MRI. Thus, the establishment of MRI as a standard medical imaging tool in the 1980s led to a revolution in the ability to see inside the human body.

1.2 The emergence of cognitive neuroscience

Our fascination with how the brain and mind are related is about as old as humanity itself. Until the development of neuroimaging methods, the only way to understand how mental function is organized in the brain was to examine the brains of individuals who had suffered damage due to stroke, infection, or injury. It was through these kinds of studies that many early discoveries were made about the localization of mental functions in the brain (though many of these have come into question

subsequently). However, progress was limited by the many difficulties that arise in studying brain-damaged patients (Shallice, 1988).

In order to better understand how mental functions relate to brain processes in the normal state, researchers needed a way to image brain function while individuals performed mental tasks designed to manipulate specific mental processes. In the 1980s several groups of researchers (principally at Washington University in St. Louis and the Karolinska Institute in Sweden) began to use positron emission tomography (PET) to ask these questions. PET measures the breakdown of radioactive materials within the body. By using radioactive tracers that are attached to biologically important molecules (such as water or glucose), it can measure aspects of brain function such as blood flow or glucose metabolism. PET showed that it was possible to localize mental functions in the brain, providing the first glimpses into the neural organization of cognition in normal individuals (e.g., Posner et al., 1988). However, the use of PET was limited due to safety concerns about radiation exposure, and due to the scarce availability of PET systems.

fMRI provided exactly the tool that cognitive neuroscience was looking for. First, it was safe, which meant that it could be used in a broad range of individuals, who could be scanned repeatedly many times if necessary. It could also be used with children, who could not take part in PET studies unless the scan was medically necessary. Second, by the 1990s MRI systems had proliferated, such that nearly every medical center had at least one scanner and often several. Because fMRI could be performed on many standard MRI scanners (and today on nearly all of them), it was accessible to many more researchers than PET had been. Finally, fMRI had some important technical benefits over PET. In particular, its spatial resolution (i.e., its ability to resolve small structures) was vastly better than PET. In addition, whereas PET required scans lasting at least a minute, with fMRI it was possible to examine events happening much more quickly. Cognitive neuroscientists around the world quickly jumped on the bandwagon, and thus the growth spurt of fMRI began.

1.3 A brief history of fMRI analysis

When the first fMRI researchers collected their data in the early 1990s, they also had to create the tools to analyze the data, as there was no "off-the-shelf" software for analysis of fMRI data. The first experimental designs and analytic approaches were inspired by analysis of blood flow data using PET. In PET blood flow studies, acquisition of each image takes at least one minute, and a single task is repeated for the entire acquisition. The individual images are then compared using simple statistical procedures such as a t-test between task and resting images. Inspired by this approach, early studies created activation maps by simply subtracting the average activation during one task from activation during another. For example, in the study by Kwong et al. (1992), blocks of visual stimulation were alternated with blocks of no stimulation. As shown in Figure 1.3, the changes in signal in the visual cortex

1.3 A brief history of fMRI analysis

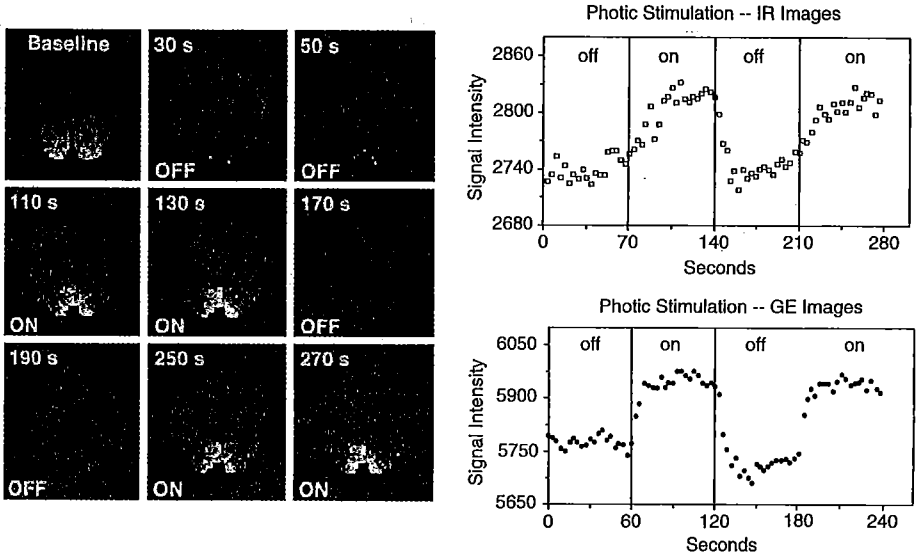

Figure 1.3. Early fMRI images from Kwong et al. (1992). The left panel shows a set of images starting with the baseline image (top left), and followed by subtraction images taken at different points during either visual stimulation or rest. The right panel shows the timecourse of a region of interest in visual cortex, showing signal increases that occur during periods of visual stimulation.

were evident even from inspection of single subtraction images. In order to obtain statistical evidence for this effect, the images acquired during the stimulation blocks were compared to the images from the no-stimulation blocks using a simple paired t-test. This approach provided an easy way to find activation, but its limitations quickly became evident. First, it required long blocks of stimulation (similar to PET scans) in order to allow the signal to reach a steady state. Although feasible, this approach in essence wasted the increased temporal resolution available from fMRI data. Second, the simple t-test approach did not take into account the complex temporal structure of fMRI data, which violated the assumptions of the statistics.

Researchers soon realized that the greater temporal resolution of fMRI relative to PET permitted the use of event-related (ER) designs, where the individual impact of relatively brief individual stimuli could be assessed. The first such studies used trials that were spaced very widely in time (in order to allow the hemodynamic response to return to baseline) and averaged the responses across a time window centered around each trial (Buckner et al., 1996). However, the limitations of such slow event-related designs were quickly evident; in particular, it required a great amount of scan time to collect relatively few trials. The modeling of trials that occurred more rapidly in time required a more fundamental understanding of the BOLD hemodynamic response (HRF). A set of foundational studies (Boynton et al., 1996; Vazquez & Noll, 1998; Dale & Buckner, 1997) established the range of event-related fMRI designs for which

the BOLD response behaved as a linear time invariant system, which was roughly for events separated by at least 2 seconds. The linearity of the BOLD is a crucial result, dramatically simplifying the analysis by allowing the use of the General Linear Model and also allowing the study of the statistical efficiency of various fMRI designs. For example, using linearity Dale (1999) and Josephs & Henson (1999) demonstrated that block designs were optimally sensitive to differences between conditions, but careful arrangement of the events could provide the best possible ER design.

The noise in BOLD data also was a challenge, particularly with regard to the extreme low frequency variation referred to as "drift." Early work systematically examined the sources and nature of this noise and characterized it as a combination of physiological effects and scanner instabilities (Smith et al., 1999; Zarahn et al., 1997; Aguirre et al., 1997), though the sources of drift remain somewhat poorly understood. The drift was modeled by a combination of filters or nuisance regressors, or using temporal autocorrelation models (Woolrich et al., 2001). Similar to PET, global variation in the BOLD signal was observed that was unrelated to the task, and there were debates as to whether global fMRI signal intensity should be regressed out, scaled-away, or ignored (Aguirre et al., 1997).

In PET, little distinction was made between intrasubject and group analyses, and the repeated measures correlation that arises from multiple (at most 12) scans from a subject was ignored. With fMRI, there are hundreds of scans for each individual. An early approach was to simply concatenate the time series for all individuals in a study and perform the analysis across all timepoints, ignoring the fact that these are repeated measures obtained across different individuals. This produced "fixed effects" inferences in which a single subject could drive significant results in a group analysis. The SPM group (Holmes & Friston, 1999) proposed a simple approach to "mixed effects" modeling, whose inferences would generalize to the sampled population. Their approach involved obtaining a separate effect estimate per subject at each voxel and then combining these at a second level to test for effects across subjects. Though still widely in use today, this approach did not account for differences in intrasubject variability. An improved approach was proposed by the FMRIB Software Library (FSL) group (Woolrich et al., 2004b; Beckmann & Smith, 2004) that used both the individual subject effect images and the corresponding standard error images. Although the latter approach provides greater sensitivity when there are dramatic differences in variability between subjects, recent work has shown that these approaches do not differ much in typical single-group analyses (Mumford & Nichols, 2009).

Since 2000, a new approach to fMRI analysis has become increasingly common, which attempts to analyze the information present in patterns of activity rather than the response at individual voxels. Known variously as multi-voxel pattern analysis (MVPA), pattern information analysis, or machine learning, these methods attempt to determine the degree to which different conditions (such as different stimulus classes) can be distinguished on the basis of fMRI activation patterns, and also

to understand what kind of information is present in those patterns. A particular innovation of this set of methods is that they focus on making predictions about new data, rather than simply describing the patterns that exist in a particular data set.

1.4 Major components of fMRI analysis

The analysis of fMRI data is made complex by a number of factors. First, the data are liable to a number of artifacts, such as those caused by head movement. Second, there are a number of sources of variability in the data, including variability between individuals and variability across time within individuals. Third, the dimensionality of the data is very large, which causes a number of challenges in comparison to the small datasets that many scientists are accustomed to working with. The major components of fMRI analysis are meant to deal with each of these problems. They include

- **Quality control:** Ensuring that the data are not corrupted by artifacts.
- **Distortion correction:** The correction of spatial distortions that often occur in fMRI images.
- **Motion correction:** The realignment of scans across time to correct for head motion.
- **Slice timing correction:** The correction of differences in timing across different slices in the image.
- **Spatial normalization:** The alignment of data from different individuals into a common spatial framework so that their data can be combined for a group analysis.
- **Spatial smoothing:** The intentional blurring of the data in order to reduce noise.
- **Temporal filtering:** The filtering of the data in time to remove low-frequency noise.
- **Statistical modeling:** The fitting of a statistical model to the data in order to estimate the response to a task or stimulus.
- **Statistical inference:** The estimation of statistical significance of the results, correcting for the large number of statistical tests performed across the brain.
- **Visualization:** Visualization of the results and estimation of effect sizes.

The goal of this book is to outline the procedures involved in each of these steps.

1.5 Software packages for fMRI analysis

In the early days of fMRI, nearly every lab had its own home-grown software package for data analysis, and there was little consistency between the procedures across different labs. As fMRI matured, several of these in-house software packages began to be distributed to other laboatories, and over time several of them came to be distributed as full-fledged analysis suites, able to perform all aspects of analysis of an fMRI study.