

The Challenge of Functional Magnetic Resonance Imaging

William F. Eddy*
Mark Fitzgerald†
Christopher Genovese‡
Audris Mockus§
Department of Statistics
Carnegie Mellon University
Pittsburgh, PA 15213-3890.

January 4, 1996

1 Introduction

Functional Magnetic Resonance Imaging (fMRI) is an extremely promising and rapidly developing technique used by cognitive neuropsychologists to obtain images of the *active* human brain. Images are obtained while the subject is engaged in a set of cognitive tasks designed to isolate specific brain functions, and the psychologists attempt to use the observed patterns of neural activation to understand and localize these functions. In contrast to traditional methods for mapping brain function, fMRI is non-invasive and allows the study of high-level cognitive processes such as language, visual attention, and problem solving. Since fMRI involves no known toxicity, each subject can be imaged many times, which improves precision and facilitates more sophisticated analyses. As such, fMRI promises to play a vital role in discerning the functional organization of the brain.

2 Functional Magnetic Resonance Imaging (fMRI)

Recent developments in magnetic resonance imaging (MRI) have greatly increased the speed with which images of the human brain can be formed and this makes it suitable

*Professor of Statistics. Partially supported by ONR Contract N00014-91-J-1024, and NSF Grants IBN-9418982 and DMS-9505007.

†Graduate Student. Partially supported by NIH Grant MH15758.

‡Assistant Professor. Partially supported by NSF Grant DMS-9505007.

§Member of Technical Staff, Bell Laboratories. Partially supported by the Center for the Neural Basis of Cognition.

for studying brain functions. An MR scanner is a multi-million dollar device which, by subjecting its contents to carefully modulated magnetic fields and recording the resulting radio signal, produces the Fourier transform of the magnetic field spin density for a particular atomic isotope. Then computing the inverse Fourier transform of the digitized signal reveals an image of the (magnetic field spin density of the) contents of the scanner.

Without going into the detailed physics and neurobiology that relate the magnetic field to brain activity, suffice it to say that increased neuronal activity induces an increase in blood flow to the region of activity (to deliver glucose to the neurons). This increased flow results in an increase of oxygenated blood in the small veins that drain the active region because the increased activity does not require much extra oxygen. The more oxygen carried by the hemoglobin in the blood the smaller the magnetic field generated by the iron in the hemoglobin (the oxygen acts as a magnetic shield) and consequently the less interference with the local magnetic field generated by, e.g., hydrogen nuclei (protons). By mid-1991 researchers had demonstrated that MRI can detect the changes in blood oxygenation caused by brain function and consequently the technique is known as fMRI. Among the first studies to use MRI to assess functional neural activity in humans are [1, 2, 3]. The latter two introduced the now common Blood Oxygenation Level Dependent (BOLD) technique just described for characterizing activation in the brain.

There are several important features of fMRI compared to other imaging techniques. First, the signal comes directly from functionally induced changes. Second, it provides both functional and anatomical information. Third the spatial resolution is on the order of 1 or 2 millimeters. Fourth, there is little known risk from fMRI. Finally, the change in signal due to brain activity is quite small (on the order of 1%) and, in particular, smaller than the noise (on the order of 2%). This last feature means that, utilizing current technology, it is necessary to average a large number of images in order to detect the regions of activation.

3 A Typical Experiment

The simplest fMRI experiment entails the performance of two cognitive tasks which differ in some specific detail. A number of images are gathered during each task and averaged within task. The difference between the average images for the two tasks provides information about the location in the brain of the cognitive function represented by the difference of the two tasks.

An actual experiment might proceed as follows. A subject lies in the MRI magnet with a head restraint intended to minimize movement. A set of preliminary anatomical images are studied to determine the location within the brain where the repeated functional images will be taken. The subject practices each of the two tasks for about a minute each, responding to the task, for example, by pushing a button with the right thumb. The subject performs one of the tasks repeatedly while images are recorded and then switches to the other task. In some of our smaller experiments we are recording 100 images for each task. In order to eliminate left-right effects the entire experiment is repeated with the subject using the left thumb to respond. Thus there are a total of 400 images in this simple experiment. It takes the scanner less than 20 minutes to acquire this amount of data.

The acquired images are multi-slice images with, typically, seven slices; each slice is

128x128 voxels with voxel dimensions roughly 2mm x 2mm x 7mm. Because most of the work we have done to date has been on the two-dimensional slices of these images we will henceforth think in terms of the $7 \times 400 = 2800$ individual slices.

4 Data Processing

The processing of the 2800 slices from this small experiment in order to detect the regions of activation is a massive task. The raw data is 128x128x2800 32-bit words which occupies 256MB of disk storage. Simply moving this amount of data around is a time-consuming task. Currently, the available bandwidth between the MR scanner and the workstation where we perform the processing is under 200K bytes per second; thus it requires nearly 30 minutes to simply move the data for this small experiment to our workstation. There are plans in place to substantially increase the bandwidth by the end of this calendar year.

Currently, the actual data processing is roughly as follows. We begin with an adjustment to account for inhomogeneity in the main static magnetic field. (We expect, in the future, to implement a further adjustment to account for non-linearity in the secondary dynamic magnetic field.) Then, we perform a “baseline” adjustment to correct for miscalibration of the analog-to-digital converter. (We expect, in the future, to implement a further “jitter” adjustment to account for very small errors in the timing of the data acquisition.) Then we perform a “mean” adjustment to correct for uncontrolled drift in the signal strength. (We expect, in the future, to implement a further pixelwise detrending to account for local drift within the image.) Then we perform an “outlier” adjustment to correct for shot noise. (We expect, in the future, to implement more sophisticated methods for addressing the fact that the data do not follow a Gaussian distribution.) We refer to the data at this point in the processing as the *corrected* data.

Unfortunately, because of the length (in time) of an experiment, the subject will almost certainly move. We address that problem both through the use of a head clamp and through a motion-correction procedure. We calculate the inverse Fourier transform to produce an image for the purposes of estimating the motion. Our motion correction procedure is complicated: first, by a non-linear optimization technique we estimate the amount of movement required to align each image and, second, we adjust the corrected data to account for this movement. We then calculate the inverse Fourier transform of the corrected and motion-corrected data to produce the actual image. At this point we are ready to begin what is called the “statistical” analysis of the data. The average is computed within slices within tasks and then the difference between tasks within slice is calculated. Finally, a statistical test is performed on each of the resulting differences to determine the regions of activation. Depending on the computing power of the workstation performing the calculations and depending on the precise details of the calculations, this can take anywhere from several days down to about twelve hours of processing time.

Our data processing software is designed as a processing pipeline of separate programs. This has the great advantage that modules can be easily interchanged if this will benefit the results. Also, new intermediate steps can be easily inserted at any stage. There is some disadvantage in that the act of storing intermediate results can consume considerable time. Nonetheless, we feel quite strongly that keeping the processing highly modularized is very

beneficial because of the flexibility it provides.

5 Statistical Challenges

The statistical challenges in the analysis of fMRI data are difficult and manifold. They all revolve around our understanding the nature of the noise and its effect on successfully detecting regions of activation. There are two general approaches to dealing with the noise in fMRI experiments. The first is to try to remove the source of the noise; we pursue this approach aggressively. The second is to model the noise through statistical methods; we also pursue this approach aggressively. We believe that both approaches are absolutely necessary.

Noise arises from a variety of sources. A fundamental source of noise is the vibration of the atomic nuclei in the imaged material. This cannot be reduced except by lowering the temperature toward absolute zero. Unfortunately, this noise is not spatially or temporally homogeneous but depends on both the anatomical structure and the function we are trying to detect. Inhomogeneity of the magnetic field, mechanical vibration, temperature instability of the electronics, etc., are all machine-based sources of noise. The machine-maintenance technicians work to limit these sources. The details of how the magnetic field is modulated to produce an image (known as a pulse sequence) effect the noise; we are engaged in studies to assess the relationship.

Physiological processes of the body such as respiration, heartbeat, and peristalsis effect the signal in ways that, in principle, can be modeled. We have begun planning experiments to gather data which might allow us to successfully model the cardiac and respiratory cycles because our more experienced colleagues believe that this is one of the primary sources of noise. Such an experiment is going to require synchronized recording of many images and the associated cardiac and respiratory information. This will be followed by a modelling effort which will view the sequence of images as the dependent variable and the cardiac and respiratory variables as predictors. Unfortunately, there is an interaction between the pulse sequence and the noise caused by physiological processes. This effort will thus require a family of models for each pulse sequence.

Movement of the subject between images is another source of noise. The standard algorithm for image registration in functional neuroimaging, called AIR [4], works on reconstructed images. It is extremely computationally intensive: registration of images obtained from a single experiment can take as much as 24 hours of computer time. Subject movement appeared to us to be the simplest of the sources to understand and address. We have developed an alternative algorithm [6] for registering the images which operates in the Fourier domain. This method has proven to be more accurate than AIR, less prone to artifacts, and an order of magnitude more efficient. By differentially weighting regions in the Fourier domain, the method can also be made less sensitive to spurious signals that have a strong influence on image domain techniques. It is also readily generalizable to three-dimensional image registration, although we have not yet completed that work.

Finally, there is subject to subject variation. We have not yet focused on this question simply because the experimenters focus their experiments on individual subjects.

All of these sources affect our ability to detect regions of activation. When we began this work, active voxels were being detected by performing an independent t-test on each of the

16384 voxels in an image. We were approached with the question: How should we correct for the “multiple comparisons?” Bonferroni corrections do not result in any “significant” voxels. Ultimately we will have to build a complex spatial-temporal model of the images which allows us to answer the real question: Where are the active regions?

We have developed another approach [5] for identifying active regions, which is called the contiguity threshold method. The idea is to increase the reliability of identification by using the fact that real activation tends to be more clustered than artifactual activation caused by noise. Empirical evidence strongly suggests that this method provides a significant improvement in sensitivity. Of course, although it is more robust than voxel-wise tests, this method, too, depends on simplistic assumptions; we intend it as a stop-gap measure, to be eventually supplanted by more sophisticated analyses.

6 Computational Challenges

There are three important aspects of the computation. First, the amount of data from a large experiment approaches 1 GB. Any computations on a data set of this size require considerable time on a workstation. Second, there are no sensible ways to reduce the data during the earlier processing steps to speed up the processing. Third, because most of the computations are done on an image-by-image basis (or even on a slice-by-slice basis), there is a tremendous opportunity to speed things up with parallel or distributed methods.

Currently, our standard processing does not take advantage of the inherent parallelism. However, we have just begun experimenting (on our local network of workstations) with Parallel Virtual Machine (PVM) implementations of some of the most time-consuming steps in the processing. Simultaneously, we have begun plans to move the computations to a Cray T3D with 512 processors. In addition to just wanting to get things done faster, another reason for this plan is that we would like to perform the computations while the subject of the experiment is in the scanner and use the results as a guide for further experimentation during the same scanning session.

7 Discussion

We have begun a serious effort to study and improve the statistical methodology of fMRI, and we have made some important preliminary steps.

One of the most fundamental questions about fMRI experiments is the question of reproducibility. If we run the experiment a second time immediately following the first with no intervening time, how similar will the results be? If we wait for a period of time? If we remove the subject from the scanner? If we repeat the experiment next month? We have begun to address this question; our preliminary results are reported in [8].

The analysis of functional Magnetic Resonance Imaging data can in many ways be viewed a prototype for a class of statistical problems that are arising more and more frequently in applications: namely, large data sets derived from a complex process with both spatial and temporal extent. There is a wealth of opportunities for the development of new statistical methodologies, and many of these ideas will apply to a variety of problems beyond neuroimaging.

8 Acknowledgements

We are indebted to a large number of colleagues at various local institutions. We list them here alphabetically both to thank them for their contributions to our work and to indicate the absolutely immense size of an undertaking of this kind. All of these people have made genuine contributions to our work either through teaching us, collecting data for us, computer programming, providing problems to solve, or assisting us in our work in other ways. The list includes psychiatrists, psychologists, physicists, electrical engineers, computer scientists, statisticians, computer programmers, and technicians. The organizational codes are given in this footnote¹.

Marlene Behrman, Ph.D., Psychology, CMU; Carlos Betancourt, B.S., MRRC, UPMC; Fernando Boada, Ph.D., MRRC, UPMC; Todd Braver, GS, Psychology, CMU; Patricia Carpenter, Ph.D., Psychology, CMU; Betty Jean Casey, Ph.D., Psychiatry, UPMC; Sam Chang, GS, MRRC, UPMC; Jonathan Cohen, M.D., Ph.D., Psychology, CMU; WPIC, UPMC; Denise Davis, B.S., MRRC, UPMC; Michael DeCavalcante, Undergraduate, CMU; Steven Forman, M.D., Ph.D., VAMC; WPIC, UPMC; Joseph Gillen, B.S., MRRC, UPMC; Nigel Goddard, Ph.D., PSC; Mark Hahn, B.S., LRDC, PITT; Murali Haran, Undergraduate, CMU; Marcel Just, Ph.D., Psychology, CMU; Caroline Kanet, Undergraduate, CMU; Timothy Keller, Ph.D., Psychology, CMU; Paul Kinahan, Ph.D., PETRC, UPMC; Benjamin McCurtain, B.S., WPIC, UPMC; Robert Moore, M.D., Ph.D., WPIC, UPMC; Thomas Nichols, B.S., GS, Statistics, CMU; PETRC, UPMC; Douglas Noll, Ph.D., MRRC, UPMC; Leigh Nystrom, Ph.D., Psychology, CMU; Jennifer O'Brien, B.S., GS, MRRC, UPMC; Brock Organ, B.S., Psychology, CMU; Robert Orr, B.S., Psychology, CMU; Julie Price, Ph.D., PETRC, UPMC; David Rosenberg, Ph.D., WPIC, UPMC; Walter Schneider, Ph.D., LRDC, PITT; David Servan-Schreiber, Ph.D., WPIC, UPMC; Steven Small, Ph.D., WPIC, UPMC; John Sweeney, Ph.D., WPIC, UPMC; Talin Tasciyan, Ph.D., MRRC, UPMC; Keith Thulborn, M.D., Ph.D., MRRC, UPMC; David Townsend, Ph.D., PETRC, UPMC; James Voyvodic, Ph.D., MRRC, UPMC; Richard Zemel, Ph.D., Psychology, CMU.

¹CMU = Carnegie Mellon University; GS = Graduate Student; LRDC = Learning Research and Development Center; MRRC = Magnetic Resonance Research Center; PETRC = Positron Emission Tomography Research Center; PITT = University of Pittsburgh; PSC = Pittsburgh Supercomputer Center; UPMC = University of Pittsburgh Medical Center; VAMC = Veterans Administration Medical Center; WPIC = Western Psychiatric Institute and Clinic.

References

- [1] Belliveau, J.W., Kennedy, D.N., McKinstry, R.C., Buchbinder, B.R., Weisskoff, R.M., Cohen, M.S., Vevea, J.M., Brady, T.J., and Rosen, B.R. (1991). "Functional mapping of the human visual cortex by magnetic resonance imaging," *Science*, **254**, 716-719.
- [2] Kwong, K.K., Belliveau, J.W., Chesler, D.A., Goldberg, I.E., Weisskoff, R.M., Poncelet, B.P., Kennedy, D.N., Hoppel, B.E., Cohen, M.S., Turner, R., Cheng, H., Brady, T.J., and Rosen, B.R. (1992). "Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation," *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 5675.
- [3] Ogawa, S., Tank, D.W., Menon, D.W., Ellermann, J.M., Kim, S., Merkle, H., and Ugurbil, K. (1992). "Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping using MRI," *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 5951-5955.
- [4] Woods, R., Cherry, S. and Mazziotta, J. (1992). "Rapid automated algorithm for aligning and reslicing PET images," *Journal of Computer Assisted Tomography*, **16**, 620-633.
- [5] Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., and Noll D.C. (1995). "Improved assessment of significant change in functional magnetic resonance imaging (fMRI): Use of a cluster size threshold," *Magnetic Resonance in Medicine*, **33**, 636-647.
- [6] Eddy, W.F., Fitzgerald, M., and Noll D.C. (1995). "Fourier domain registration of MR images." Submitted.
- [7] Bandettini, P.A., Jesmanowicz, A., Wong, E.C., and Hyde, J.S. (1993). "Processing strategies for time-course data sets in functional MRI of the human brain," *Magnetic Resonance in Medicine*, **30**, 161.
- [8] Eddy, W.F., Behrmann, M., Carpenter, P.A., Chang, S.Y., Gillen, J.S., Just, M.A., Keller, T.A., Mockus, A., Tasciyan, T.A., and Thulborn, K.R. (1995). "Test-Retest reproducibility during fMRI studies: Primary visual and cognitive paradigms. *Proceedings of the Society for Magnetic Resonance, Third Scientific Meeting*, 843.