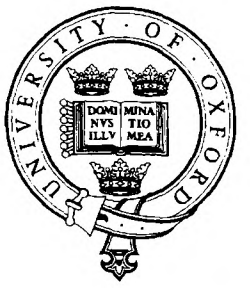
Motion Correction for Functional Magnetic Resonance Images

Peter R Bannister



Michaelmas Term, 2003

Robotics Research Group

Department of Engineering Science

University of Oxford

This thesis is submitted to the Department of Engineering Science, University of Oxford, in fulfilment of the requirements for the degree of Doctor of Philosophy. This thesis is entirely my own work and, except where otherwise stated, describes my own research.

Peter R. Bannister, Green College

Motion Correction for Functional Magnetic Resonance Images

Peter R Bannister

Green College
University of Oxford

Abstract

This work addresses the distortions in Functional Magnetic Resonance Images (FMRI) caused by subject motion. FMRI is a non-invasive technique which shows great promise in providing researchers and clinicians with neurological information both about healthy subjects and clinical patients by mapping functional activation within the brain using Echo Planar Imaging (EPI). If reliable information is to be obtained from these images, motion correction must be carried out in order to remove or suppress the artefacts arising from subject movement. This work begins by using exploratory data techniques to describe these artefacts so that they can be characterised according to their origin and spatio-temporal manifestation.

Based on testing of the accuracy and consistency of existing rigid-body motion correction methods on FMRI data, a new registration algorithm — Motion Correction using the FMRIB Linear Image Registration Tool (MCFLIRT) — has been developed. It is shown that while MCFLIRT is both more accurate and more robust than previous methods, rigid-body registration schemes in general cannot completely remove the distortions associated with motion and so subsequent analysis of the images may still be inaccurate. Furthermore, it is demonstrated that failure to use a sufficiently detailed model of subject motion in FMRI can in fact lead to degradation of the images through the use of existing motion correction algorithms. Based on these findings, alternative schemes including non-rigid registration and adaptive real-time methods are evaluated.

Leading on from this investigation, a framework for Temporally-Integrated Geometric EPI Realignment (TIGER), incorporating both spatial and temporal information about the images, is proposed. An implementation based on this novel modality-specific model is developed and tested against existing rigid-body registration methods. Results show that this new approach is able to achieve significantly more accurate results than previous methods. The quality of correction provided by this new approach brings more subtle artefacts in the data to the fore, suggesting a number of avenues of further research in this area. These are outlined in the final chapter of the thesis.

Thesis submitted for the degree of Doctor of Philosophy
at the University of Oxford

Michaelmas Term, 2003

Acknowledgements

My thanks go first and foremost to my supervisors: I am very fortunate to be able to acknowledge Professor Sir Mike Brady as a continual source of wisdom, enthusiasm and common-sense, while Dr Mark Jenkinson was not only unfailingly patient, but also generous with his time, his math, his debugging skills, his spare room and his car. I would also like to highlight the contribution made by my diligent and insightful examiners, Dr Daniel Rueckert (Imperial College) and Dr Paul Newman, whose comments have contributed greatly to the final form of this thesis.

Members of FMRIB and the Robotics Research Group provided numerous ideas and key discussions over the past four years, in particular Christian Beckmann, Stuart Clare, Stephen Smith, Mark Woolrich and Yongyue Zhang. I'm also indebted to those who 'volunteered' to proof-read this thesis: my supervisors, my mother, Luis Freire, Mark Woolrich and Paul Wyatt. Grateful thanks go to: everyone at the Howard Florey Institute at the University of Melbourne, where a productive two months were spent in the Neuroinformatics Lab; the fellows and staff of Green College, notably Sir John Hanson; Dr Ian Reid, not least for providing me with an income over the last year; Professor David Murray, for inspiring my interest in computer vision and encouraging me to pursue the research that has culminated in this thesis; Dr Paul Beardsley, for fuelling my enthusiasm for research while I worked at MERL; Professor Lionel Tarassenko and my colleagues in the Signal Processing and Neural Networks Research Group, for being consistently accommodating, supportive and encouraging during the final stages of my doctoral studies.

Last, but certainly not least, I should like to say a big thank you to my friends in Oxford for keeping me (relatively) balanced over the past few years (with apologies in advance to those I have failed to mention individually): Nephthys 2001 and the Bad Boy; Pat Lockley and ECBC; the Worcester and St Anne's 1995 intake; Spanners and Test-Tubes; Ace How and Ugly Bob – thanks for always reminding me how much less fun than academia the world of work is, and for keeping my gaming repertoire up to date; Rob – cheers for getting me out on my bike when writing up was threatening to turn me into a permanent hermit; and

of course Mom, DJ and Floss – thank you so much for your love and support. Finally, Dad and Nes: you have each, by your example, guided and inspired me, especially with regard to this work.

The research summarised in this thesis was funded by an EPSRC Doctoral Studentship Award. Financial support is also gratefully acknowledged from the Universities Superannuation Scheme, the Medical Research Council, the International Society for Magnetic Resonance in Medicine, the Department of Engineering Science, the Oxford Medical Research Fund and Green College.

Related Publications

Journal Papers

Advances in Functional and Structural MR Image Analysis and Implementation as FSL
Stephen M. Smith, Mark Jenkinson, **Peter R. Bannister**, Christian F. Beckmann, Timothy E.J. Behrens, Marilena De Luca, Ivana Drobnjak, David Flitney, Heidi Johansen-Berg, Rami K. Niazy, James Saunders, John Vickers, Mark W. Woolrich, Yongyue Zhang, J. Michael Brady, Paul M. Matthews
Neuroimage *to appear.*

Variability in FMRI: A Re-Examination of Intersession Differences
S.M. Smith, C.F. Beckmann, N. Ramnani, M.W. Woolrich, **P.R. Bannister**, M. Jenkinson, P.M. Matthews, and D.J. McGonigle.
Human Brain Mapping *in press.*

Improved Optimisation for the Robust and Accurate Linear Registration and Motion Correction of Brain Images
M. Jenkinson, **P.R. Bannister**, J.M. Brady, and S.M. Smith.
Neuroimage 17(2):825–41.

Conference Proceedings

Flaws in Existing Spatio-Temporal and Temporo-Spatial Realignment of FMRI Data
P.R. Bannister, J.M. Brady and M. Jenkinson.
Medical Image Understanding and Analysis, September 23–24, 2004, London, U.K.

Non-rigid Motion Estimation and Spatio-temporal Realignment in FMRI
P.R. Bannister, J.M. Brady and M. Jenkinson.
3rd IEEE Workshop on Articulated and Non-rigid Motion, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 27–July 2, 2004, Washington D.C., U.S.A.

TIGER A New Model For Spatio-temporal Realignment Of FMRI Data
P.R. Bannister, J.M. Brady and M. Jenkinson.
Workshop on Computer Vision Approaches to Medical Image Analysis, 8th European Conference on Computer Vision, May 10–15, 2004, Prague, Czech Republic.

Intersession Variability in FMRI and the Effect of Different Analysis Methods
S.M. Smith, C.F. Beckmann, N. Ramnani, M.W. Woolrich, **P.R. Bannister**, M. Jenkinson, P.M. Matthews, and D.J. McGonigle.
In the proceedings of the 9th Annual Meeting of the Organization for Human Brain Mapping Conference on Functional Mapping of the Human Brain, June 18–23, 2003, New York, U.S.A.

Spatio-temporal Realignment of FMRI Data
P.R. Bannister, S.M. Smith, J.M. Brady and M. Jenkinson.
In the proceedings of the 8th Annual Meeting of the Organization for Human Brain Mapping Conference on Functional Mapping of the Human Brain, June 2–6, 2002, Sendai, Japan.

FEAT - New Developments
S.M. Smith, M.W. Woolrich, **P.R. Bannister**, C.F. Beckmann, T.E. Behrens, J.M. Brady, S. Clare, M. De Luca, D. Flitney, D. Greve, and M. Jenkinson.
In the proceedings of the 8th Annual Meeting of the Organization for Human Brain Mapping Conference on Functional Mapping of the Human Brain, June 2–6, 2002, Sendai, Japan.

Motion Artefact Decorrelation in FMRI Analysis using ICA
P.R. Bannister, C.F. Beckmann, and M. Jenkinson.
In the proceedings of the International Society for Magnetic Resonance in Medicine 10th Scientific Meeting, May 18–24, 2002, Honolulu, Hawai'i.

Robust Affine Motion Correction in fMRI Timeseries
P.R. Bannister and M. Jenkinson.
In the proceedings of the 7th Annual Meeting of the Organization for Human Brain Mapping Conference on Functional Mapping of the Human Brain, June 10–14, 2001, Brighton, U.K.

Exploratory Motion Analysis in fMRI using ICA
P.R. Bannister, C.F. Beckmann, and M. Jenkinson.
In the proceedings of the 7th Annual Meeting of the Organization for Human Brain Mapping Conference on Functional Mapping of the Human Brain, June 10–14, 2001, Brighton, U.K.

FSL: New Tools for Functional and Structural Brain Image Analysis
S.M. Smith, **P.R. Bannister**, C.F. Beckmann, J.M. Brady, S. Clare, D. Flitney, P. Hansen, M. Jenkinson, D. Leibovici, B. Ripley, M.W. Woolrich, and Y. Zhang. In the proceedings of the 7th Annual Meeting of the Organization for Human Brain Mapping Conference on Functional Mapping of the Human Brain, June 10–14, 2001, Brighton, U.K.

Notation Summary

AIR	Automated Image Registration software package
BOLD	Blood Oxygen Level Dependency
CBF	Cerebral Blood Flow
CBV	Cerebral Blood Volume
CR	Correlation Ratio
CSF	Cerebrospinal Fluid
CT	Computed Tomography
DOF	Degrees Of Freedom
DTI	Diffusion Tensor Image
EEG	Electroencephalography
EPI	Echo Planar Image
EVI	Echo Volumetric Image
FFT	Fast Fourier Transform
FLIRT	FMRIB's Linear Image Registration Tool
fMRI/ FMRI	Functional Magnetic Resonance Image
FMRIB	Oxford Centre for Functional Magnetic Resonance Imaging of the Brain
FORCE	FMRIB's Optimized Retrospective Correction Environment software
FOV	Field Of View
FRE	Fiducial Registration Error
FSL	FMRIB's Software Library
FT	Fourier Transform
GLM	General Linear Model
IC	Independent Component
ICA	Independent Component Analysis
ICP	Iterated Closest Point
LS	Least Squares

MARE	Median Average Residual Error
MCFLIRT	Motion Correction using FLIRT
MEG	Magnetoencephalography
MELODIC	Multivariate Exploratory Linear Optimized Decomposition into Independent Components
MI	Mutual Information
MR	Magnetic Resonance
MRI	Magnetic Resonance Image
NC	Normalised Correlation
NMI	Normalised Mutual Information
NMR	Nuclear Magnetic Resonance
ONE	Orbital Navigator Echo
PD	Proton Density
PET	Positron Emission Tomography
RF	Radio Frequency
RIU	Ratio of Intensity Uniformity
RMS	Root Mean Square
RV	Random Variable
SNR	Signal-to-Noise Ratio
SPM	Statistical Parametric Map *also* Statistical Parametric Mapping analysis software
SPM99	1999 SPM software release
TIGER	Temporally Integrated Geometric EPI Realignment
TR	Time to Repeat

Contents

List of Figures

Chapter 1

Introduction

The work in this thesis addresses the issue of motion correction in Functional Magnetic Resonance Imaging (FMRI). FMRI is a relatively recent development of Magnetic Resonance Imaging (MRI) and is a non-invasive technique which can be used to form images of neural activation. The technique can answer questions about the way in which the brain works and can also characterise deficiencies due to illness or injury. The modality is based on measurements of the magnetic behaviour associated with blood flow change due to metabolic activity which can be observed using an MRI scanner (figure 1.1).

Despite these key applications a major drawback of FMRI is that it is currently possible to form only relatively low-resolution images, such as the example shown in figure 1.2, which exhibit poor signal-to-noise ratios. This makes it much harder to reliably detect activation and dynamic behaviour in a subject's brain. And, because this detection is based on the statistical comparison of many scans of the same sections of the brain, the situation is further confounded by the adverse effect of subject motion at even very low levels of a few millimetres or less. This situation is exacerbated by the fact that clinical patients will generally move far more than cooperative volunteers while in the scanner.

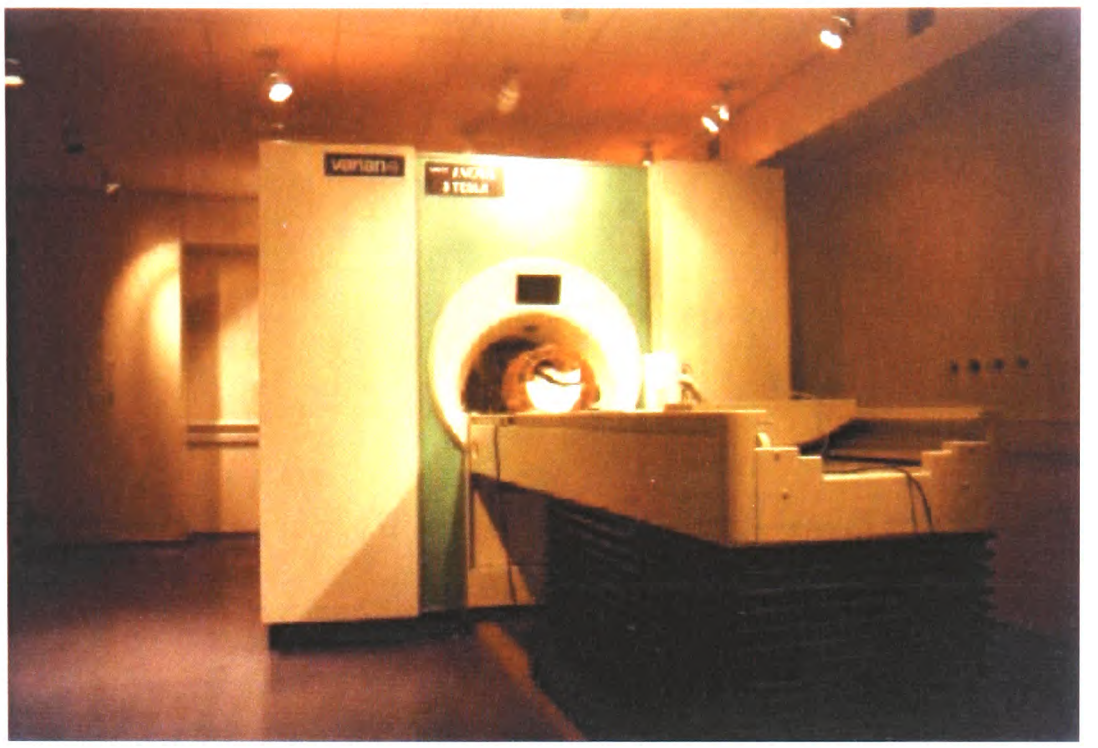


Figure 1.1: The 3T Varian/Siemens magnetic resonance imaging system installed at the Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB) where the research described in this work was carried out. The scanner is designed specifically for ultra-fast imaging for brain functional magnetic resonance imaging studies and is equipped both with a body gradient coil and a fast head gradient coil insert. In addition, it has capabilities for perfusion, diffusion and spectroscopy studies.

Furthermore, typical FMRI experiments produce data covering several hundred time-points, each characterised by a volume containing over $60,000$ 3-D pixels, or *voxels*. Given the large volume of data which must be considered from a single FMRI study and the potential uncertainty in the data caused by motion artefacts, there is some reluctance from the neuroscience community to adopt FMRI as a clinical tool.

In order for FMRI to be considered a reliable clinical tool it is essential that techniques which can compensate for this motion are developed so that consistently robust conclusions can be drawn from the data. This chapter introduces the methods and applications of FMRI and explains that while these images are potentially very informative to researchers and clinicians, they typically contain

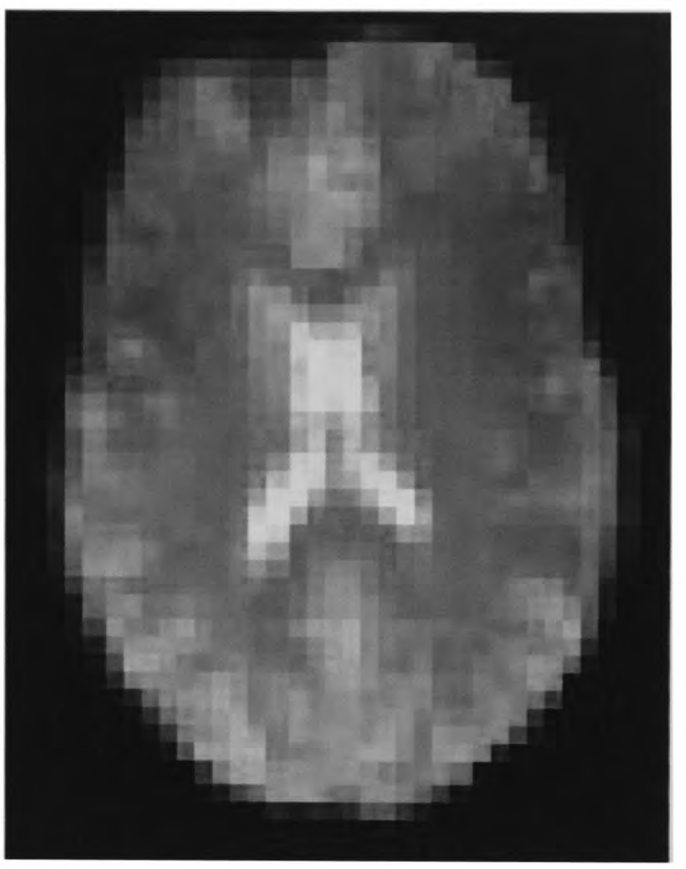


Figure 1.2: An example echo planar image demonstrates the low resolution available from the modality. The ventricles can be identified as the white cross in the centre of the brain but the delineation between grey and white matter is unclear elsewhere and certainly nowhere near the millimetre-resolution required for neuroscientific analysis.

motion-induced artefacts which impact significantly on the accuracy of the data in the images.

The aim of this thesis is to further understanding of the nature of these artefacts and to develop methods which are able to correct the data. A more detailed overview of the structure of the thesis is presented in section 1.5.

1.1 FMRI Overview

Unlike static MRI techniques, FMRI facilitates the formation of images that indicate neural activation indirectly by examining the changing oxygenation levels of blood flowing in the brain, thought to result from metabolic changes brought

about by mental activity [80]. These images serve a variety of purposes, ranging from pre-surgical location of critical functional regions [38] to furthering the understanding and treatment both of pathological conditions (such as MS or stroke) and more subtle psychological or cognitive disorders. In addition to these applications, FMRI has established itself as a powerful tool in the study of memory, perception and motor control [82]. FMRI images may also be used to augment clinical information in cases in which the results of anatomical investigation are uncertain [48] and the technique has formed part of an augmented reality system for real-time image-guided surgery [34].

1.1.1 BOLD mechanism

In order to image neural activity, it is necessary to devise methods which are able to detect the physiological manifestation of brain activation. Changes in neuronal activity are accompanied by changes in cerebral blood flow (CBF), cerebral blood volume (CBV) and blood oxygenation [151] (see figure 1.3), allowing FMRI techniques based on imaging sequences sensitive to these physiological changes to be used.

Upon activation,[1] oxygen is extracted by the cells, thereby increasing the level of deoxyhaemoglobin in the blood. This is compensated for by an increase in blood flow in the vicinity of the active cells within 10 seconds [97] or less, leading to a net increase in oxyhaemoglobin. This Blood Oxygen Level Dependency (BOLD) response is the underlying mechanism which links FMRI techniques to the study of mental activity.

Since the relative levels of oxyhaemoglobin and deoxyhaemoglobin are correlated with neural activity, it is necessary to seek some way of measuring them. One

[1] Characterised as a change in neuronal firing rates averaged over a macroscopic area.

obvious candidate for an indicator is the different colours of oxyhaemoglobin and deoxyhaemoglobin, and this has been used in optical imaging techniques [37]. There are, however, several restrictions associated with optical methods: light cannot easily penetrate the skull, and though fibre optics can be used, this would involve an invasive procedure. And, because infra-red light has relatively little energy, it is possible to detect only neurological changes close to the surface of the brain. An alternative is to exploit the difference in paramagnetic properties between oxyhaemoglobin and deoxyhaemoglobin and this is the basis of FMRI.

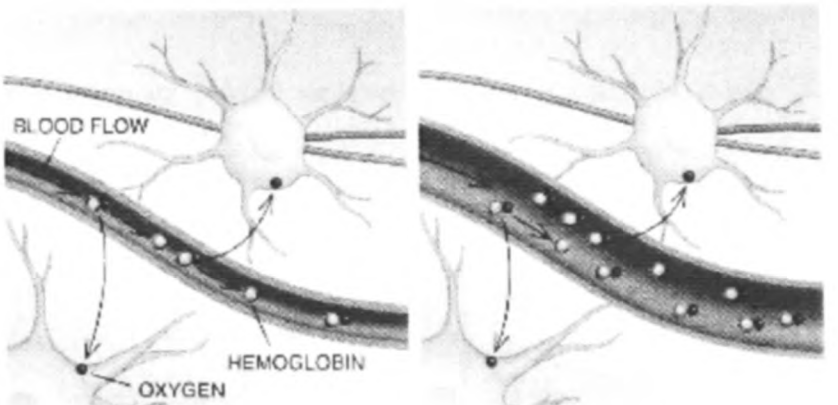


Figure 1.3: A diagram of the increased flow of oxygenated blood due to neural activation [122], known as the Blood Oxygen Level Dependency (BOLD). Increased mental activation, shown on the right, gives rise to a greater flow of haemoglobin, leading to a greater concentration of ^{1}H which in turn results in a stronger signal measured by systems (such as MRI scanners) which are sensitive to the concentration of protons in biological tissue.

1.2 The Physics of FMRI

The behaviour of nuclei in MRI is governed by quantum physics but it is sometimes easier to gain a simplified overview using classical physics. For a fuller explanation, the reader is directed toward [144] or [20].

In their resting state, the protons (^{1}H) present in biological systems can be thought of as tiny spinning 'bar-magnets', all initially randomly oriented. When

a sufficiently strong static magnetic field, B_0, is applied in the direction defined as the z-direction to the medium containing these nuclei they all align either with or against the field and precess about the z-axis at a characteristic frequency (the Larmor frequency). The Larmor frequency is governed by the strength of the magnetic field and a property of each nuclear species, the *gyromagnetic ratio* γ (figure 1.4). Each nuclear species has its own associated Larmor Frequency given as

$$\omega = \gamma B_0 \tag{1.1}$$

Typically for common NMR active nuclei, γ ranges from 3.078 MHz/T for ^{14}N to 42.575 MHz/T for ^{1}H.

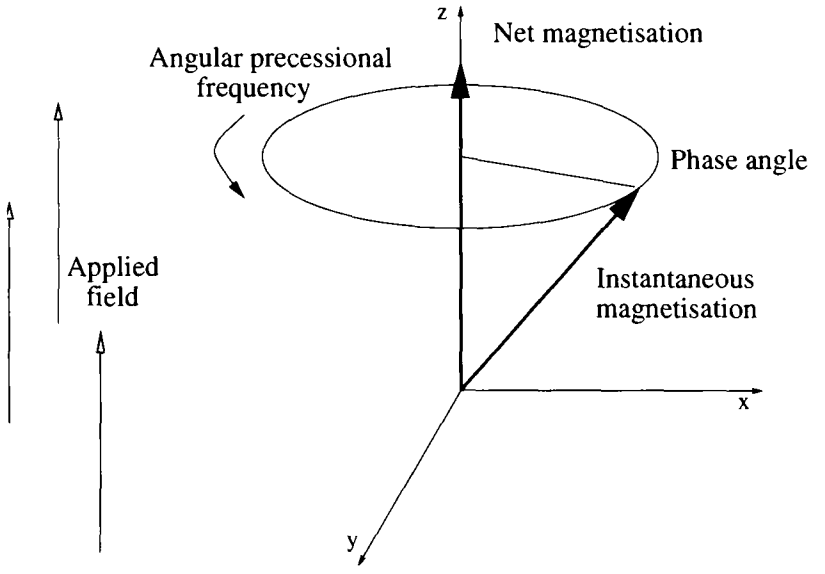


Figure 1.4: The classical model of nuclear magnetisation for an object subjected to a net magnetisation field.

This is a resonant phenomenon directly relating the resonant frequency of the nucleus to the applied field. The sum over all the nuclear magnetic moments in an object gives the *bulk magnetisation* for the object M. If an RF pulse near the resonant Larmor Frequency is then applied orthogonal to B_0, M is forced away

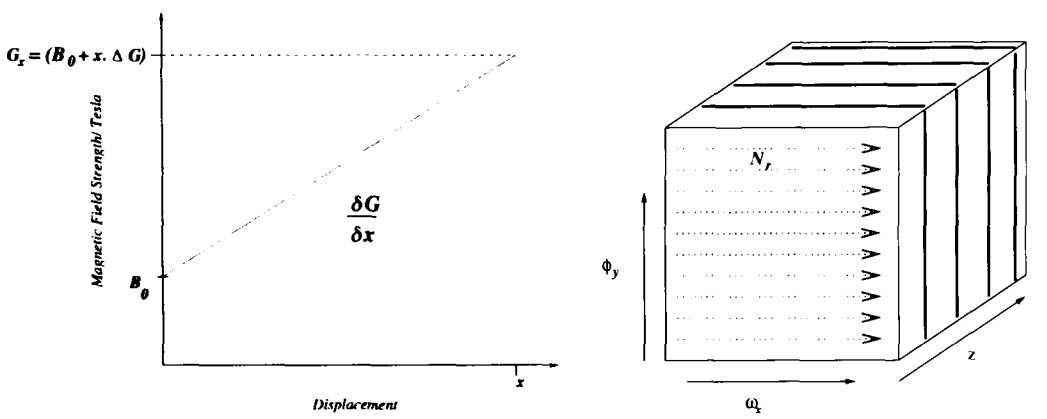


Figure 1.5: *Left:* Spatial location of a feature in 1-D using a gradient field construction. Terms correspond to equation 1.2. *Right:* Representation of a three-dimensional volume using k-space where ϕ_y, ω_x, and z denote the phase, frequency and slice encode directions respectively and N_r is the number of repeats.

from its initial alignment in the z-direction into the (x, y) plane. A weak signal, again oscillating at the Larmor Frequency, can be detected after the pulse has finished.

1.2.1 Image Formation

Given the electromagnetic phenomena described above, the question then arises: how does one form an image of a biological system by exploiting these magnetic properties? As early as the 1940s, Nuclear Magnetic Spectroscopy had established the presence of distinct grey and white matter within the brain [52] but a gradient-based system is required to locate the matter spatially along an axis in 1-D. By changing the applied field, it is possible to alter the received frequency.

In the construction (figure 1.5 *left*), an applied gradient field G_x results in a change in resonant frequency given as

$$\omega_x = \gamma(B_0 + x.G_x). \tag{1.2}$$

The frequency spectrum now detected reflects a range of distinct positions. In

order to form a two-dimensional image using NMR, two gradient coils are placed orthogonal to one another. The first stage of image formation is usually to form a 2-D representation of a 3-D object. Slice selection is a technique to isolate a single plane in the object being imaged, by exciting only the spins in that plane. To do this, an RF pulse which excites a limited frequency range is applied in the presence of a linear field gradient along the direction along which the slice is to be selected. This results in the excitation of only those nuclei whose Larmor Frequency, which is dictated by their position, is within the band of frequencies of the applied RF pulse.

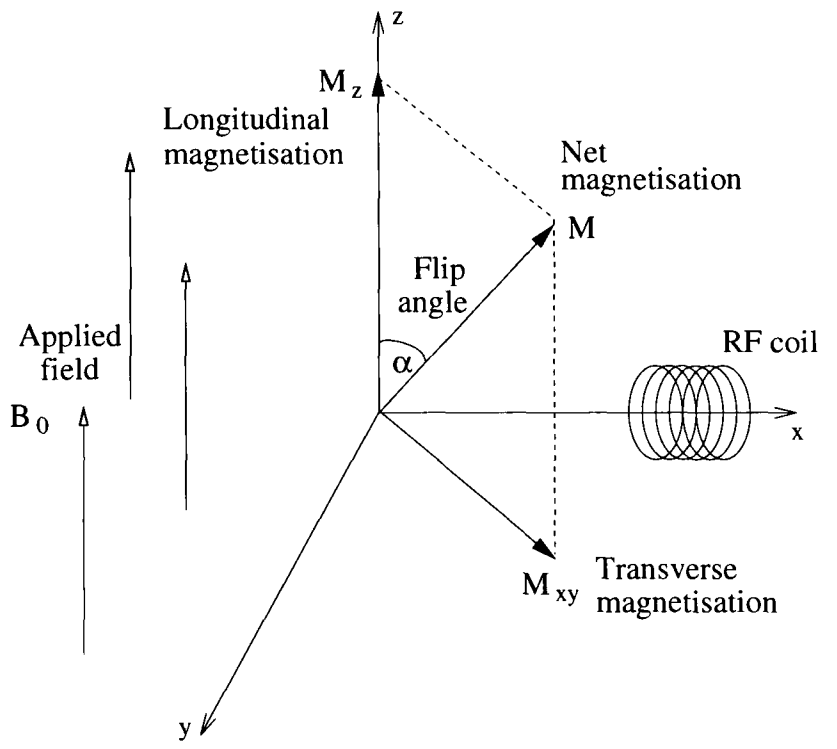


Figure 1.6: The arrangement of the 2-D slice excitation coil with respect to the applied field B_0. The transverse magnetisation M_{xy} can be determined as $M \sin\alpha$ and the longitudinal magnetisation M_z is given as $M \cos\alpha$.

The received signal is proportional to the transverse magnetisation (shown as M_{xy} in figure 1.6) which has an associated 'flip angle', α (figure 1.4). By applying gradients which modulate the B_0 field in the x- and y- directions, the relative phases of the nuclei at each point in the slice can be altered. The total magnetic

field (in the z-direction) at a point $\mathbf{x} = (x, y)$ is given as

$$B(\mathbf{x}, t) = B_0 + x.G_x(t) + y.G_y(t). \tag{1.3}$$

The instantaneous phase change is related to the instantaneous B field by the Larmor equation, so the phase at time t is given by the integral

$$\phi(\mathbf{x}, t) = \int \gamma B(\mathbf{x}, t) \mathrm{d}t. \tag{1.4}$$

The emitted signal can be written as

$$S(\mathbf{x}, t) = A(\mathbf{x}).e^{j\phi(\mathbf{x},t)} \tag{1.5}$$

where $A(\mathbf{x})$ is the amplitude of the transverse magnetisation at $\mathbf{x}$, assuming that it is approximately constant over the acquisition period. The total received signal is obtained by integrating over the whole object

$$S(t) = \int S(\mathbf{x}, t) \mathrm{d}\mathbf{x} = \int A(\mathbf{x}).e^{j\phi(\mathbf{x},t)} \mathrm{d}\mathbf{x}. \tag{1.6}$$

This can be associated with the Fourier Transform of $A(\mathbf{x})$ which relates physical space to *k-space* [121], since

$$A(\mathbf{k}) = \int A(\mathbf{x}).e^{-j\mathbf{k}.\mathbf{x}} \mathrm{d}\mathbf{x} \tag{1.7}$$

and comparing terms with equation 1.6 gives

$$\mathbf{k}.\mathbf{x} = k_x.x + k_y.y = -\gamma \int B(\mathbf{x}, t)\mathrm{d}t \tag{1.8}$$

$$= -\gamma \left(x \int G_x(t)\mathrm{d}t + y \int G_y(t)\mathrm{d}t \right) \tag{1.9}$$

where the constant $B_0 t$ term has been dropped as it is modulated out by the receiver circuitry. Therefore, by using appropriate gradient waveforms, the transformed signal $A(\mathbf{k})$ can be measured for any particular point in k-space.

1.2.2 Applications

To date, MRI has mostly been used to picture relatively static anatomy with exquisite results [54] like those shown in figure 1.7. Using these images, information can be extracted relating to a wide range of conditions. An example of this is the measurement of asymmetry between the left and right hemispheres of the brain. An overly high measure of symmetry is believed to correlate with schizophrenia [106]. MRI techniques have also been developed to allow the study of moving anatomy such as the heart and liver.

Recently, the development of fast imaging techniques have made possible the study of neural activation. In this situation, the dynamics the acquisition attempts to capture vary more rapidly than the time required to acquire an image using current static imaging techniques.

The acquisition speed can be increased by using Echo Planar Imaging (EPI), based on raster scanning of k-space. In this case, a single RF excitation samples both phase-encode direction ϕ_y and frequency-encode direction ω_x. While this achieves the necessary acquisition speeds for individual slices which are grouped along the z-axis to form volumes (typically a volume must be acquired every 3

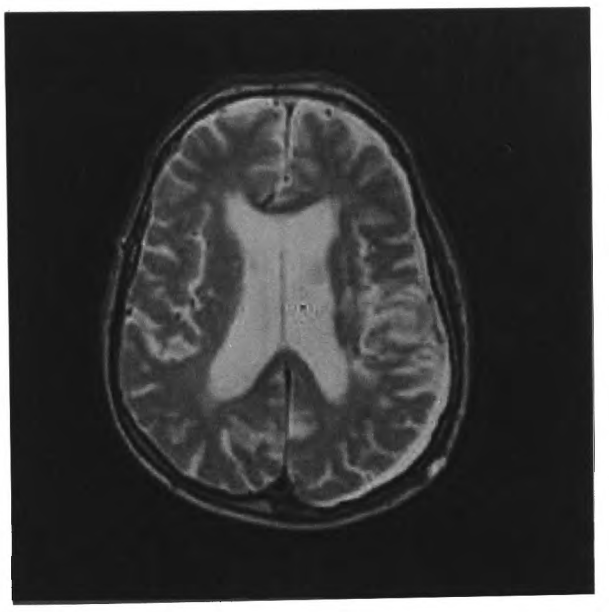


Figure 1.7: A structural MR image demonstrating the high level of detail available for static imaging modalities. Contrast this with the low-resolution EPI shown in figure 1.2.

seconds so that the temporal dynamics of the haemodynamic response can be captured), the signal-to-noise ratio (SNR) suffers and as a consequence image resolution must be lowered.

1.3 Modelling Neurological Response

It is important to understand the nature of the signal which subsequent analysis seeks to recover, in view of the poor SNR expected in FMRI. Figure 1.8 shows a typical unfiltered temporal sequence for a voxel within the visual cortex.

The data were acquired during an experiment consisting of alternating sets of 10 scans first with, and then without, visual stimulation (applied using LED goggles). This type of experiment can be characterised by its *paradigm* which describes the different states of the stimulus and the duration for which it is applied in each case. The paradigm is depicted on figure 1.8 as a square wave

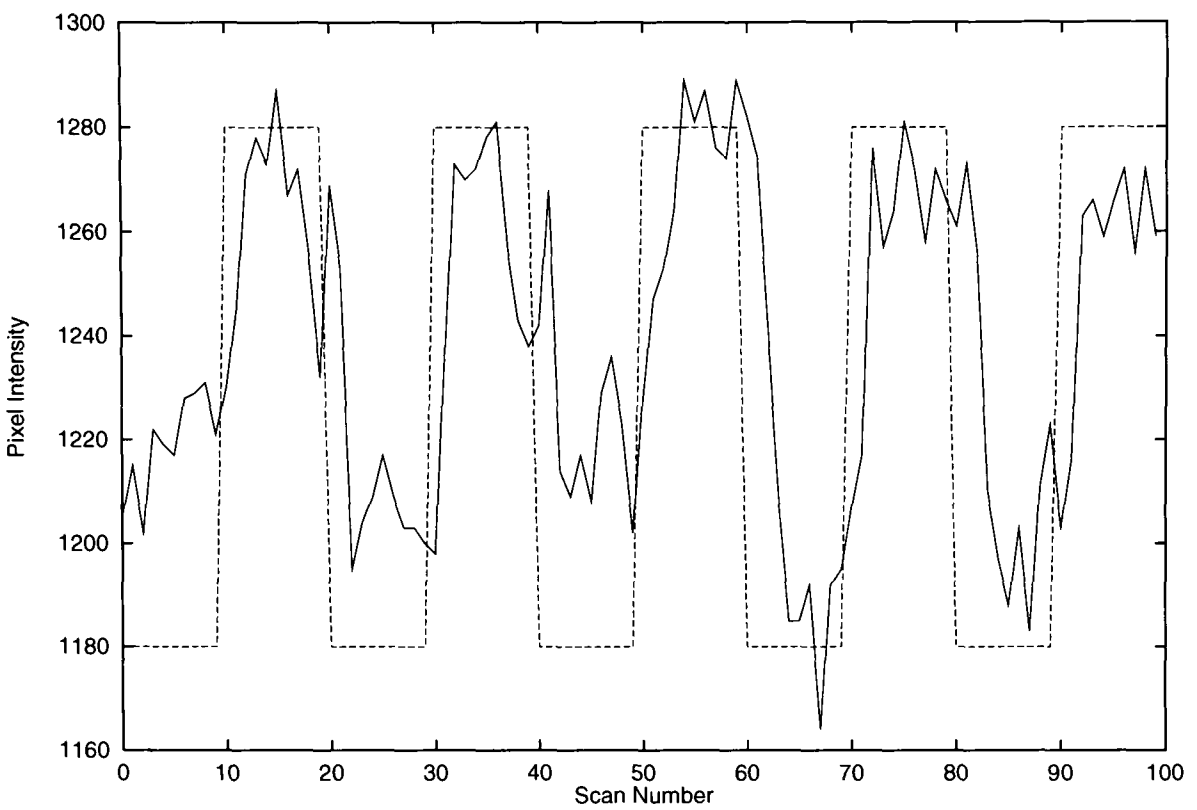


Figure 1.8: A typical 1-D temporal series (solid line) corresponding to a periodic stimulation (dashed line) of a single voxel within the visual cortex taken from an experiment featuring a strong visual stimulus. *(Data courtesy Michael Brammer, Institute of Psychiatry, King's College London)*

with a cyclic period of 20 scans[2] often referred to as a *boxcar design.*

In such a paradigm there are two 'populations' of temporal signal responses, corresponding to time-points where the stimulus is applied or is not applied. Asserting that a voxel 'responds' to the stimulus amounts to demonstrating that there is a significant difference between these two populations. The simplest way to test this hypothesis is by a Student's T-test [115].

Note that in figure 1.8 there is an apparent lag between the change of stimulus and the neurological response. A number of models for this response have been

[2]The amplitude of the square wave is arbitrary and is not scaled to represent the stimulus.

proposed in order to classify activation response including the common form shown in equation 1.10. It is also evident in this series that even in an area of supposedly high activation there is a large amount of confounding noise.

1.3.1 Detecting Activation

One means of testing the null hypothesis[3] is to use a general linear model (GLM). By using the GLM on separate time-series, it is possible to achieve a mass univariate approach. Let $y(t)$ be the time-series in question. A measure of its similarity to a predicted response (for example, BOLD) is required. The response is typically represented by the convolution of a model of the BOLD haemodynamic response function $h(t)$ with a neuronal response function $c(t)$ which represents some contrast parameter (the stimulus function). The predicted time-series, $x(t)$ is given as

$$x(t) = \int c(t-\tau).h(\tau)\, d\tau \tag{1.10}$$

where the impulse response function $h(\tau)$ is often modelled as a Gamma function:

$$h(\tau) = (\lambda^{\tau} e^{-\lambda})/\tau! \tag{1.11}$$

The 2-D histogram of the z-statistics[4] for a particular slice of data, written SPM$\{z\}$, will, under the null hypothesis, have an approximately Gaussian distribution [114] of unit variance and zero mean.

The model, which can include more than one experimental stimulus or *ex-*

[3]The term 'null hypothesis' is used in FMRI analysis to refer to the default assumption that no activation is present in the data.

[4]In FMRI, 'z-statistic' refers to the value computed over all the time-points for each voxel by comparing the area under tail of the T-test distribution with that of the normal distribution.

planatory variable (EV), is given for a single voxel time-series of length T as

$$\mathbf{Y} = \mathbf{XB} + \mathbf{e} \tag{1.12}$$

where $\mathbf{Y}$ is a $T \times 1$ vector containing the observed data at the voxel, and $\mathbf{X}$ is a $T \times N_B$ matrix of N_B 'regressors' (often referred to as the design matrix). In FMRI this typically contains the assumed stimuli responses for the experiment being analysed. The $N_B \times 1$ vector $\mathbf{B}$ comprises the set of parameters to be estimated. It is assumed that the $T \times 1$ error vector $\mathbf{e}$ is multivariate Normally distributed:

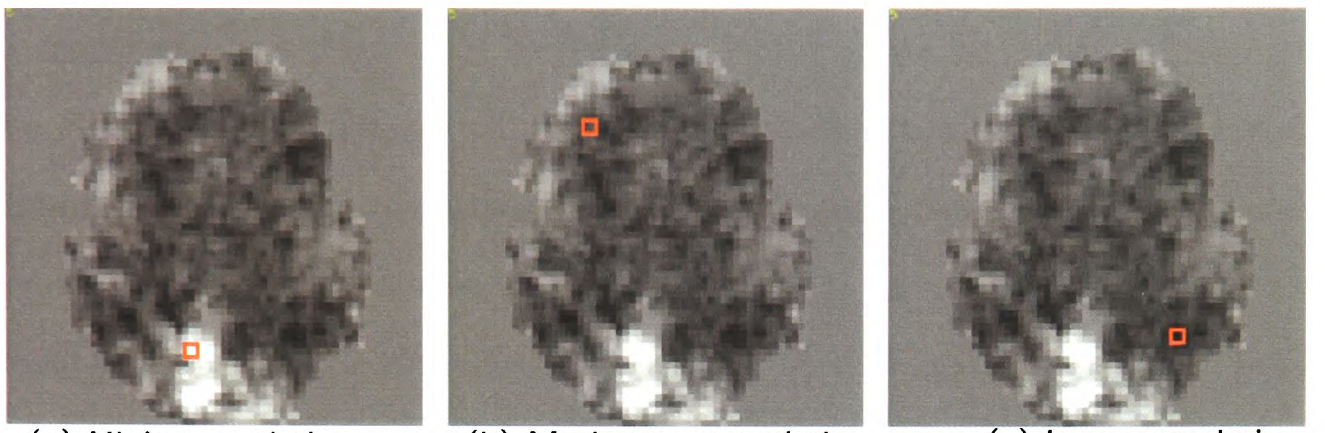
$$\mathbf{e} \sim MVN(0, \sigma^2\mathbf{V}) \tag{1.13}$$

where $\mathbf{V}$ is the $T \times T$ autocorrelation matrix and σ^2 is the variance. More details regarding the estimation of these parameters can be found in Appendix I.

The significance of a large z-statistic like the one highlighted in figure 1.9(a), which corresponds to the temporal series shown in figure 1.8, is that it implies a high level of correlation between the FMRI time-series and the ideal BOLD response to the input stimulus [45], something unlikely to be attributable to random chance and therefore highly unlikely under the null hypothesis. A low z-statistic (figure 1.9(c)) corresponds to an unstimulated voxel.

Once pre-processing (which includes steps such as high-pass filtering, slice-timing correction and motion correction) has taken place, it is then possible to test for activation. Having obtained a z-statistic value for each voxel, the value is then plotted for each voxel on a 3-D map known as a Statistical Parametric Map (SPM). A slice of the SPM obtained for the experiment described in section 1.3 are shown in figure 1.9 with each 2-D map showing a different characteristic value

corresponding to areas of high, moderate and low activation respectively.



(a) High z-statistic (b) Moderate z-statistic (c) Low z-statistic

Figure 1.9: Three z-statistics from an unpaired T-test applied to data from the experiment shown in section 1.3 are marked in red. Each corresponds to a different characteristic level of activation with white indicating a strong correlation with the experimental design and hence a strong likelihood that the voxel in question responds to the mental task of interest. Conversely, a dark voxel suggests that the corresponding region of the brain is not associated with the experimental task.

1.3.2 Noise in FMRI

Both intrinsic sources of disturbance[5] and exogenous disturbances (such as intrinsic acoustic scanner noise, sampling errors and truncation) can account for up to 10% of the intensity signal [91] which may be as much as double the intensity of the activation which detection seeks to identify, although a more typical figure for the total noise would be between 0.5 and 2%. These confounding intensity fluctuations are sometimes referred to as *baseline*.

In the next section it is shown that such artefacts, especially motion during the acquisition of functional data, can significantly corrupt the images produced and therefore lead to potentially inaccurate inferences about neuronal activity.

[5]These include non-stimulus-related activation, movement-related effects, cardiac pulse, respiratory changes, ghost images, flow artefacts in the vicinity of large vessels and regional sensitivity losses (signal voids) due to susceptibility differences at tissue borders, for example the bone-brain transition.

1.4 Spatial Artefacts in FMRI

As mentioned in section 1.3.2, FMRI acquisition is beset by a range of physical and physiological confounds [77, 79] which lead to distorted and potentially uninformative data. These confounds include imperfections in the scanning hardware and in patient-related phenomena. Some of the more common effects in the latter group include pulsatile flow effects, respiratory motion, cardiac motion and bulk motion.

Some of these artefacts may be corrected for during acquisition while others can be removed either fully or partially by post-acquisition image processing techniques. In the case of pulsatile and respiratory motion, it is common to apply gating to the imaging process so that scans are acquired in the absence of major lung or heart movement [130]. Other physical effects may not be as easy to detect and/or model and in these cases a reliable method is needed to detect, separate and characterise the artefacts.

1.4.1 Exploratory Data Analysis

While there are a range of approaches which can be adopted to detect specific artefacts in FMRI data, it is possible to perform Independent Component Analysis (ICA) on the data in order to detect a range of confounds [107, 9]. Unlike the General Linear Model which is a hypothesis-driven analytical technique, the basic goal of ICA is to solve a *Blind Source Separation* problem [55] where the FMRI data (comprising several different sources including scanner artefacts, physiological and bulk patient motion as well as experimental haemodynamic response) or *observations* are explained as linear combinations of latent component variables or *source signals*.

A more detailed explanation of ICA is contained in appendix II.

Because ICA provides both spatial maps and time-courses associated with each component, it is possible in many cases to associate particular components with known artefacts either through the temporal or spatial characteristics of that particular extracted source [5].

1.4.2 Motion Artefacts

Subject motion is present in almost all studies and ranges from the cardiac and respiratory effects discussed above along with slight patient motion to, in the most extreme cases, large bulk motion arising from a seizure during scanning. In addition, it is possible to identify physiological baseline effects by comparing the components to Independent Components (ICs) obtained from a null data study (that is, where no stimulus is presented), as shown in figure 1.10.

A typical ICA of a data-set containing (in this case artificially added) motion is shown in figure 1.11. It is unlikely that simple, direct correlations between ICA time-courses and the given motion parameters will exist because the latter are estimated with respect to a fixed set of orthogonal axes while the IC decomposition will not necessarily use the same decomposition. It is possible in this case, however, to observe a striking similarity between one of the IC time-courses and the y-translation applied to the synthetic data.

Having identified a strong temporal link between the IC and the artefact, the spatial manifestation of the confound may then be characterised. In the same way an artefact whose spatial manifestation is already known may be identified, giving new information about the associated temporal profile. In the case of bulk motion, it is possible to see an intensity increase along one edge of the spatial map accompanied by an intensity dip on the opposite side of the head.

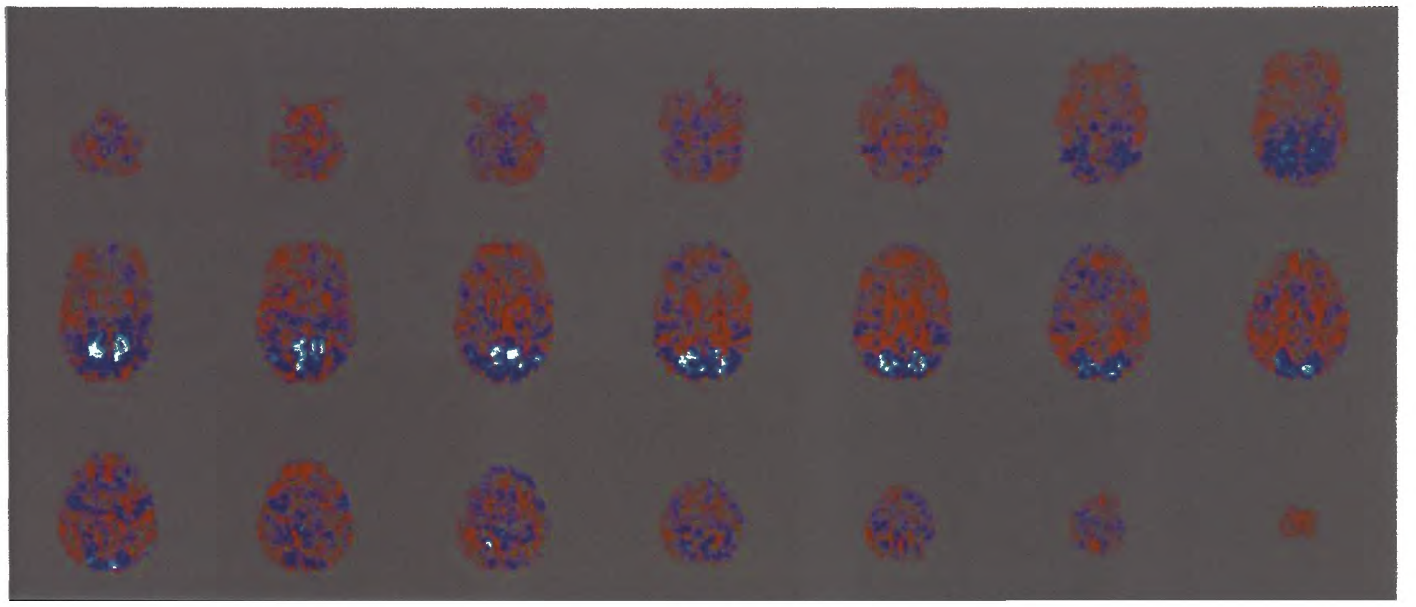


Figure 1.10: ICA spatial maps and time-courses corresponding to baseline physiological artefacts. In the top plot, the spatial maps associated with this Independent Component show that there is a large amount of spatially-localised intensity variation (coloured blue) which does not belong to the expected area of activation associated with this visual experiment (that is, the visual cortex which is located in the lower section of the brain in the middle row of spatial maps). These spatial maps are multiplied by the time-course (shown in both temporal and frequency space in the middle and bottom plots respectively where the y-axis in the former case is a dimensionless scaling factor) over the course of the experiment to give the total intensity contribution attributed to this component. The original FMRI data can be reconstructed by summing all the Independent Component spatial maps multiplied by the associated IC time-course value at each time-point. The inclusion of the Fourier transformed temporal data enables periodic signals to be more readily identified. In this example, the strong low frequency peaks in the FT plot may indicate some underlying sinusoidal noise process, further supporting the hypothesis that this IC represents the presence of baseline noise.

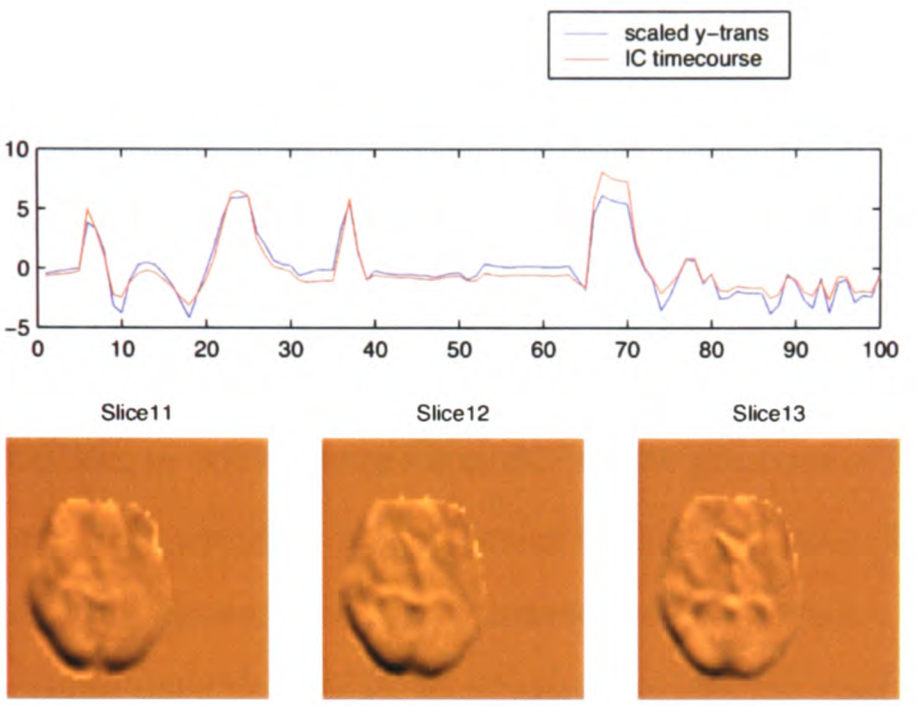


Figure 1.11: Associated time-course and spatial map from synthetic motion audiovisual FMRI data containing a strong y-translation. The FMRI data used to produce this ICA was generated by adding a series of artifical motions to a static FMRI template image to produce a motion-corrupted time-series. The x-axis of the upper plot corresponds to the time index and the y-axis measures the dimensionless amplitude, as in figure 1.10. The strong correlation between the extracted IC time-course (shown in red in the upper plot) and the original motion amplitude (scaled to overlay the IC time-course and plotted in blue) indicates that this IC represents the contribution made to the FMRI data by the motion artefact. This is confirmed by the edge-localised spatial intensity peaks (shown in black in the lower plot) which are also thought to indicate motion at the edge of the brain.

This information can assist in identifying ICs which can be argued to represent patient movement, even in cases where the temporal correspondence may not be as strong.

ICA can also be used to reveal the levels of residual motion present in data after motion correction has been applied. This is useful not only in demonstrating the effectiveness of a motion correction algorithm but also in illustrating when a simple motion correction approach will fail.

1.4.3 Acquisition Artefacts

Artefacts introduced by the acquisition process include B_0 susceptibility, slice dropout (through-slice dephasing) and $N/2$ ghosting. The presence of air–bone or air–tissue interfaces causes inhomogeneities in the static B_0 field which in turn can lead to spatial distortions (primarily in the phase-encoded direction) as shown in figure 1.12, as well as slice dropout. The latter occurs when there is a significant signal loss in one or more slices due to the presence of susceptibility-induced magnetic field gradients perpendicular to the slice plane [51], leading to a high variance in the slices affected. An example is shown in figure 1.13. Finally, $N/2$ ghosting occurs when there is an asymmetry between the even and odd lines of k-space. In this case, the data will fold along the phase-encoded direction. This can be seen in the spatial map of figure 1.14.

1.5 Motion Correction

The motivation behind this research is the need to to understand the nature of the motion artefacts common to FMRI data and to design algorithms which can apply corrections to the data in order to remove these potential confounds to the detection of activation.

While existing image registration tools have been applied to this problem with some limited success, reported in the survey in Chapter 2, it is necessary to understand the source and nature of the artefacts if they are to be corrected in order to compensate for the limited information in the data itself. This thesis presents a development of the FMRIB Linear Image Registration Tool [75] specifically for motion correction in FMRI. This is the first major contribution resulting from the research reported in this thesis.

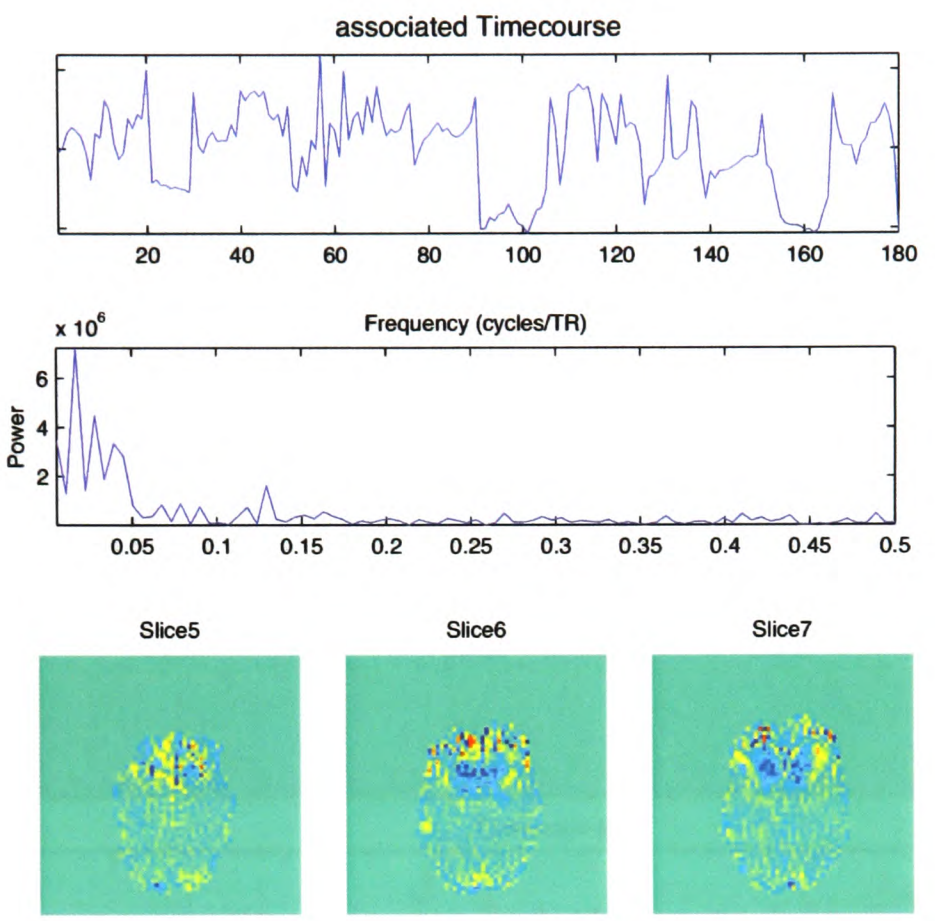


Figure 1.12: IC time-courses and selected spatial maps showing strong susceptibility artefact at the front of the brain. The upper plot represents the modulation factor at each time-point which is applied to the spatial maps in the lower plot to produce the intensity contribution made by this Independent Component. The Fourier transform of the temporal characteristic of the IC is shown in the middle plot. The IC has been identified by noting that the area of greatest intensity variation (indicated by the yellows in the spatial maps) is located at the front of the brain. This is consistent with the spatial characteristics of susceptibility-motion interactions during scanning.

Chapter 3 describes this new tool and demonstrates that it is capable of providing more accurate and robust results from motion-corrupted FMRI data than other commonly used schemes. In doing so, however, the results also highlight the shortcomings of any FMRI motion correction scheme based only on linear transformations, as a result of the presence of non-linearities in the data. In Chapter 4, a number of non-linear methods are examined and in Chapter 5 a novel technique is proposed which is tailored to the unique characteristics of

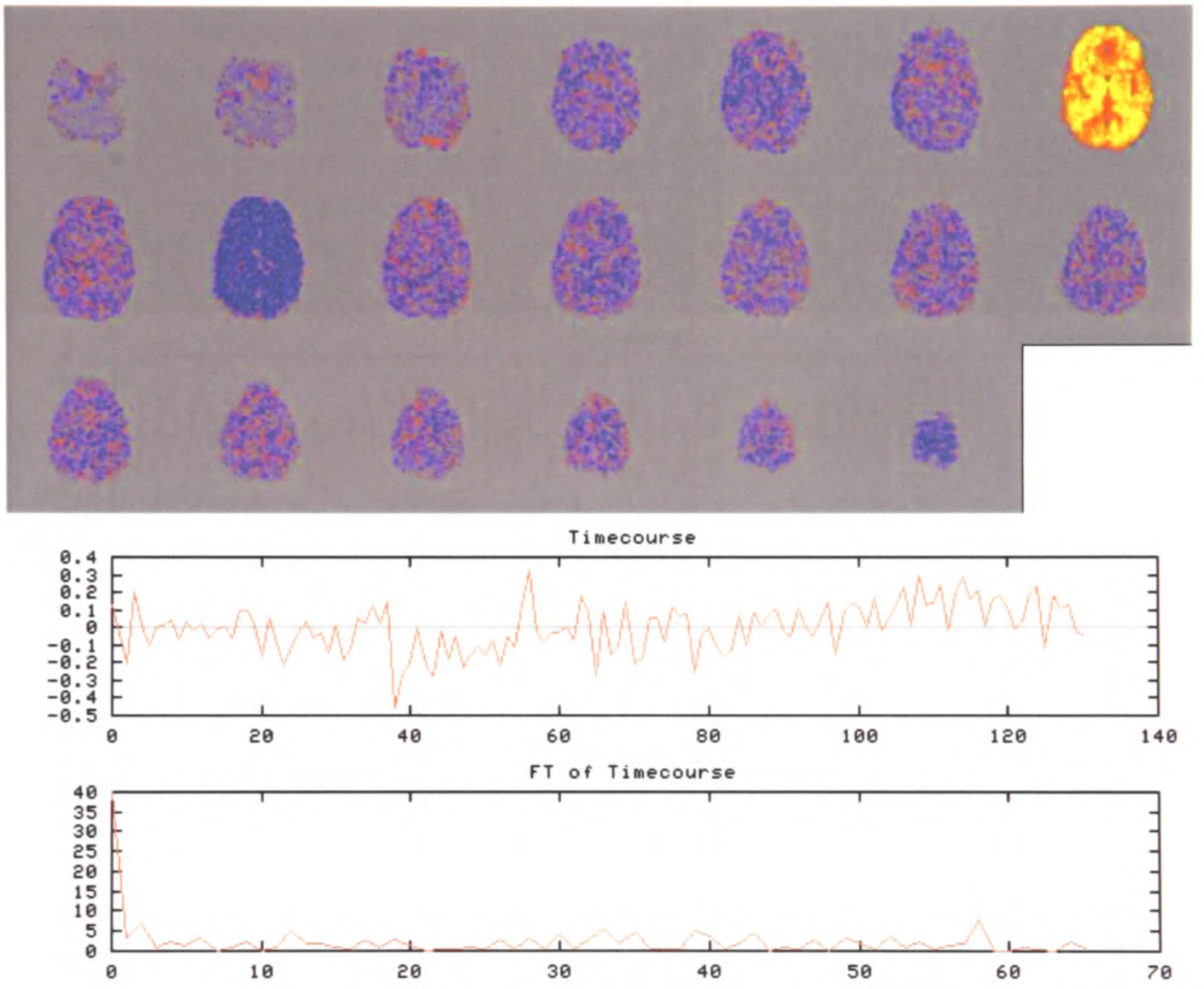


Figure 1.13: ICA spatial maps and time-courses corresponding to slice dropout. In the case of this Independent Component, the temporal plots are almost redundant as the effect of slice drop-out is clearly visible as the uniformly low signal which is visible as the yellow area in the top-right slice of the spatial map. The scanner calibration imbalance which leads to this artefact can also be seen in the second slice of the middle row of the spatial plot which indicates a complementary overall increase in intensities (indicated by the blue extreme of the colour range).

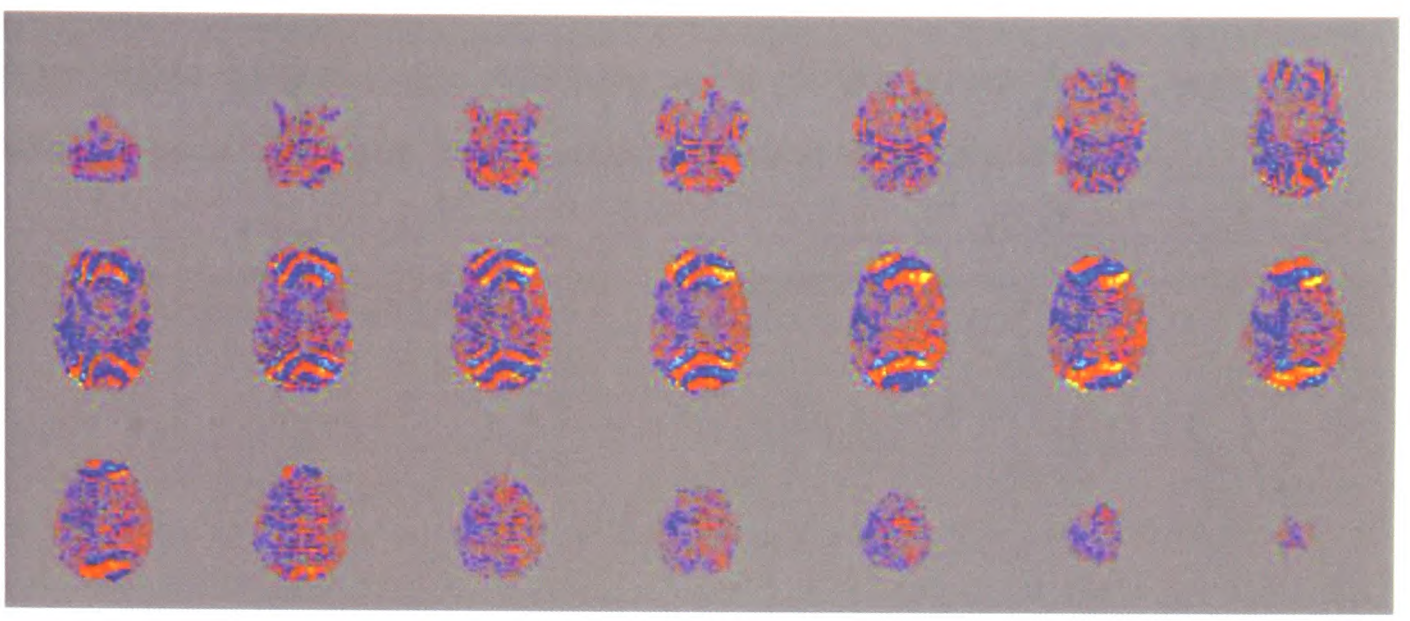


Figure 1.14: ICA spatial maps and time-courses corresponding to $N/2$ ghosting. 60Hz noise, readily identified via the lower plot of the Fourier transformed time-course of this IC, will typically indicate a scanner-related artefact. This is confirmed by the visible wraparound in the spatial maps shown in the top plot which are typical of $N/2$ or so-called 'Nyquist' ghosting.

FMRI data. This is the first reported attempt to integrate the spatial and temporal corrections required by FMRI data. A preliminary implementation, based on this new model, is developed and tested in Chapter 6.

In the final chapter of the thesis, Chapter 7, a number of issues which have not yet been addressed properly by any existing schemes are investigated and potential avenues of further research are proposed.

In Chapter 2, a review of medical imaging registration methods is presented.

These registration algorithms have been designed for a diverse range of applications within medical image processing and particular attention will be paid to those methods which have been applied to motion correction in FMRI.

Chapter 2

FMRI Motion Correction Survey

The problem of motion correction is usually considered to be a specific case of medical image registration, a process in which two or more images are aligned so that the information contained in each may be d consistently over defined physical structures or regions. In this way it is possible to pool information from a number of modalities or, as in the case of FMRI, to compare changes within a modality over time or between studies. Throughout this work, the term 'multi-modal' is used to refer to images acquired two or more different imaging protocols (for example, CT and MR) while 'mono-modal' indicates that the images under examination are assumed to have been formed using the same imaging technique.

Given this classification of the motion correction problem, a natural approach to removing motion artefacts in FMRI would appear to result from applying to the data techniques which have previously been developed for the problem of medical image registration. In this chapter, existing approaches to rigid-body registration of medical images are reviewed. Special focus is paid to those methods which have been extended to facilitate motion correction of functional time-series. The chapter concludes with an introduction to the methods used in validating the accuracy and robustness of all such registration algorithms.

While the use of affine[1] image registration techniques in FMRI has provided some useful results, the schemes discussed rely in general on a much richer level of information than FMRI is able to offer. Consequently it is necessary to develop a modality-specific approach which implements some of the ideas reviewed in the following sections but with the particular goal of FMRI motion correction.

2.1 Motivation

Subject motion is an important issue in FMRI analysis where even the slightest movement[2] during a scan can displace the voxel location corresponding to a distinct physical area from one image to the next. In addition, the change in signal intensity due to motion can be far greater than the BOLD effect itself, particularly at tissue boundaries, at the edge of the brain or near major vessels where the difference in intensity between the two adjoining areas is greatest (figure 2.1).

Before analysis of the activation patterns corresponding to different parts of the brain can take place, the acquired images must be motion-corrected in order for the physical and image coordinate systems to be coincident. This is an extension of the registration problem because conventional MR images, often generated in different modalities (e.g. T1 vs. T2, PD vs. T2), must be aligned before they can be used in further clinical studies (See figure 2.2).

While it may prove that implementing a variety of existing image registration methods (already used for tasks such as brain atlasing [25], the task of registering a patient's brain to a standardised anatomical atlas) would be sufficient to correct for motion artefacts, consideration must also be given to the fact that the

[1]Where, in addition to the 6 rigid-body parameters controlling translation and rotation in the 3 principal axis directions, each axis is permitted to scale and/or skew, giving a total of 12 degrees of freedom.

[2]1 to 2 degrees of rotation or a few millimetres translation.

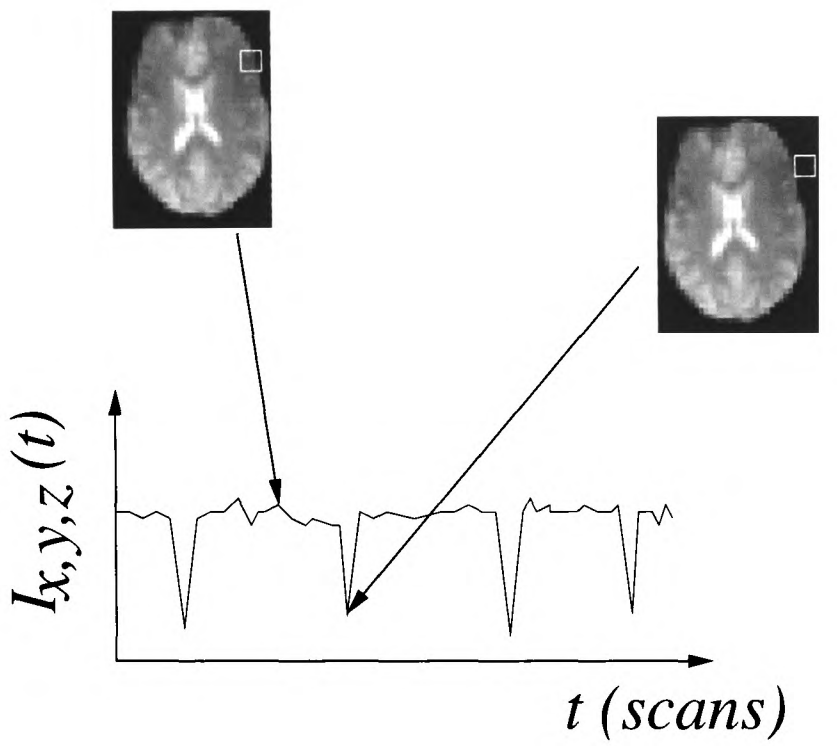


Figure 2.1: Simulation of the effect of motion on the time-course of a voxel located near an intensity boundary. The voxel marked in white corresponds to the same point in scanner coordinates in both images and its time-course is shown in the plot. Subject motion (in this case a translation to the left) has meant that intensity value of this voxel exhibits severe fluctuations between samples.

magnetisation of the nuclei is affected by the position of the patient's head within the scanner. This suggests a model-based approach which takes into account the physics of the acquisition process which does not assume either a linear (rigid or affine) or non–linear transformation. Due to the low resolution of FMRI and the fact that the skull is a rigid object, it is reasonable *a priori* to suppose that an affine or even a rigid-body registration scheme which seeks to identify a limited set of motion parameters may be feasible and sufficient for motion correction. Such schemes are discussed and compared below.

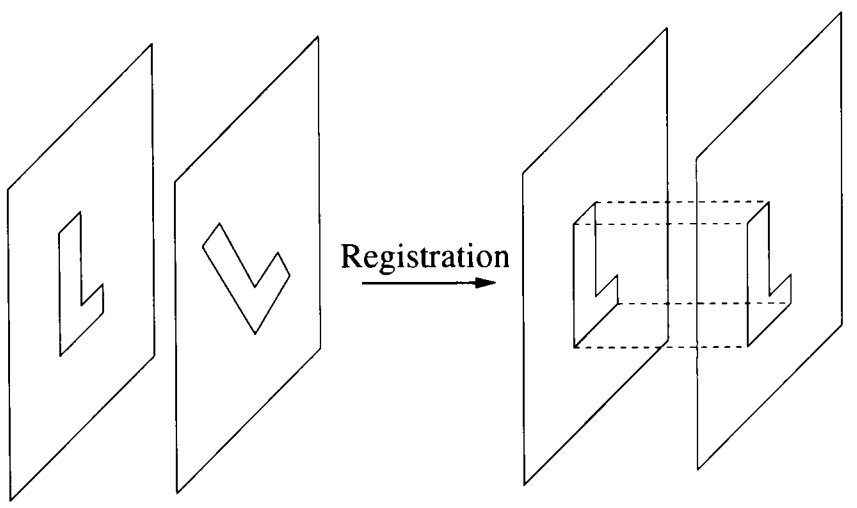


Figure 2.2: Registration transforms images to align their object view.

2.2 Registration in MRI

Within medical image registration literature there is a broad distinction made between multi-modal and mono-modal realignment. A coarse division of multi-modal applications would be into anatomical and functional-anatomical registration. In the case of mono-modal registration the images exhibit similar intensity ranges and a uniform contrast pattern across different tissue types.

Given the assumption that the images being registered are of the same patient, mono-modal registration (which includes FMRI motion correction where the registration is EPI-to-EPI) is significantly easier than multi-modal registration because of the high degree of similarity between image pairs.

2.2.1 Types of Registration

Registration procedures can always be split into three aspects: the *problem statement*, the *registration paradigm* and the *optimisation procedure* [102]. The majority of applications outside the operating theatre allow for off-line registration so real-time speed is not usually an issue.

Within the remainder of this section, each of the three aspects above is explained. Following on from this general classification, the alternative approaches available within each of these areas are discussed in more detail with special focus paid to methods which may be appropriate for (intra-modal, intra-subject) FMRI motion correction.

Problem Statement

This determines the classification of the registration problem under consideration according the modalities involved (either mono-modal or multi-modal), the subjects of the images (intra-subject, inter-subject or patient-to-atlas) and the type of object represented by the images, in this case the brain.

Registration Paradigm

This influences the nature of the registration basis which can be either intrinsic or extrinsic. Commonly used extrinsic methods (which are based on objects attached to the patient's head to aid matching) allow for fast, easy registration and since registration parameters can be computed analytically, there is no need for complex optimisation algorithms. The underlying assumption behind these methods, however, is that the fiducial markers are not only available, but also appropriate to the correction which is being inferred and readily identifiable in the images. FMRI data does not conform to any of these requirements which precludes the application of standard computer vision techniques.

Other factors which impact on the form of registration paradigm are the nature of the transformation (whether it can be efficiently described as rigid, affine, projective or curved), the level of user interaction allowed or required by

the realignment scheme[3] and finally the domain of the transformation which may be either global over the entire image or local where subsections of the image have their own transformation defined.

Generally, local transformations are curved and global transformations are either rigid or affine (that is, they map straight lines to straight lines). The latter are most commonly used where the image scaling factors are unknown or suspected to be incorrect (e.g. MR images with geometric distortions).

If a transformation maps straight lines onto curves it is called *curved*, *elastic* or *non-linear*. Most applications represent curved transformations in terms of a local vector displacement (disparity) field

$$y = x + t(x) \tag{2.1}$$

where x represents the original data, y denotes the transformed data and t is the transformation. Alternative representations can be formed as transformations in terms of the old coordinates.

The term 'local transformation' is reserved for transformations that are composites of at least two transformations determined on sub-images that cannot simply be described as a global transformation. Since local information of the anatomy is essential to provide an accurate local curved transformation, applications are often semi-automatic and excellently suited for inter-subject and atlas registration.

Maintz and Viergever [102] note that many methods spend over 90% of their computation time examining registrations at a resolution level that would hardly benefit from human intervention. If they perform robustly, such methods are

[3]The three levels of registration algorithms are: automatic, interactive or semi-interactive where the user may supply initial conditions or select one of a number of proposed hypotheses.

better left automated. Human interaction also complicates the validation of registration methods as it is not easily quantified or controlled. Of the intrinsic methods, anatomical landmark and segmentation based methods are commonly semi-automatic and the geometrical landmark and voxel intensity-based methods are usually automatic, requiring no user initialisation.

Optimisation Procedure

The choice of optimisation procedure may influence the level of interaction but it most directly affects the parameters which are computed directly by the registration scheme and those which are determined by finding an optimum on some function, usually referred to as the *cost function*, defined on the parameter space.

In the latter case, the similarity measure is, it is hoped, well behaved (quasi-convex) so that one of the standard and well-documented optimisation techniques can be used. These include Powell's method, the downhill simplex method, Brent's method, Levenberg-Marquardt optimisation, Newton-Raphson iteration (where the error between two images is written as a low-order Taylor expansion and subsequently minimised), stochastic search methods, gradient descent methods, simulated annealing, geometric hashing and quasi-exhaustive search methods. Frequently, extensions are made to the basic scheme including multi-resolution and multi-scale approaches, in order to reduce the number of transformations to be examined and to attempt to avoid local minima. Some methods employ custom non-standard optimisation methods such as the ICP algorithm [12], created for rigid-model-based registration. Many applications use more than one optimisation technique: coarse and fast followed by fine and slower.

2.2.2 Intrinsic over Extrinsic Methods

Until recently, the stereotactic frame (see figure 2.3) provided the gold standard for registration accuracy. Since it is used for navigation in neurosurgery it does not add an additional invasive strain on the patient. However, use of such a frame purely to aid registration is to be avoided.

Figure 2.3: A stereotactic frame shown attached to a mannequin. *(Image ©2002 Mountain Neurosurgery and Spine, Hendersonville, NC)*

Non-image-based registration can be achieved by calibrating coordinate systems between two scanners in the same physical space and where the patient lies motionless between experiments. In reality, this is currently only practical in MRI and Computed Tomography (CT) if motion constraints (e.g. head restraints) are used, or if a combined CT-PET machine is used, or for ultrasound systems with a spatial (optical) location device. This method is also used for registering the position of surgical tools mounted on a robot arm in relation to images.

Extrinsic methods restrict the registration transformation to be either rigid or affine, because, by definition, it is impossible to include patient-related information. Intrinsic methods are, however, considerably more versatile and can be based on a limited set of salient marker points, the alignment of segmented

binary structures (usually surfaces) or from direct measurement of voxel intensity values.

Landmark-based Methods

Within intrinsic methods, landmarks may be anatomical or geometric. The identification of landmarks is technically a segmentation procedure. Registration methods based on landmarks are mostly used for rigid or affine transformations although, in theory, if there are sufficiently many points, more complex transformations may be accommodated.

Segmentation-based registration can be rigid model-based, where extracted surfaces are used as the sole input for the alignment procedure (which itself may be non-rigid); or deformable model-based, where one image is elastically deformed to fit the second image. A drawback of segmentation-based methods is that the registration accuracy is limited to the accuracy of the segmentation step which is usually performed semi-automatically. While this thesis addresses the topic of deformable model-based registration in detail in chapter 4, a brief introduction is included here to highlight the relative advantages of rigid-body schemes in terms of simplicity and numerical behaviour.

For deformable model-based methods, the optimisation criterion is usually defined locally and incorporates constraints imposed on the deformable curve (snake or active contour [14]). The 3-D extension of the model is sometimes referred to as a net. The procedure is carried out iteratively, often using incremental deformations using localised spline functions.

Deformable model approaches deform a template model in the first image to fit either a segmented second image or a feature (e.g. an edge) in an unsegmented second image. Rigid models are used mostly for intra-subject registration while

deformable approaches are often necessary for inter-subject and atlas registration applications.

Deformable models are sensitive to the choice of initial position needed to ensure convergence (so rigid pre-registration is typically used) and the local template deformation can be erratic if the target structure differs significantly from the template structure. This means that they are more suitable for finding locally curved transformations and less so for global rigid or affine transformations. A major application area is cortical registration of MR images which is important because it enables the direct transfer of the segmentation of one cortex to another.

Intensity-based Methods

There are two distinct approaches to voxel intensity-based registration methods. One is to reduce immediately the image grey value content to a representative set of scalars and orientations; the second is to use the full image content throughout the registration process. Principal axes and moments-based methods are the prime examples of reductive registration methods. Although the result is not usually very accurate, the methods are used where fast approximate registration is appropriate. The method can also be used as a coarse pre-registration step for other methods.

Methods using the full image content can be applied in almost any medical application area, using any type of transformation. Nearly all methods considered in [102] are automatic and although some cover curved transformations in the theory, this is a relatively sparsely covered topic in the literature.

There are many suitable cost function measures commonly in use: cross-correlation [84], Fourier-domain-based cross-correlation and phase only cross-correlation [31], minimisation of grey value variation within segments [26], max-

imisation of the mutual information [18], histogram clustering [63], minimisation of joint entropy [98], maximisation of zero crossings [140], minimisation of the squared intensity differences [67], matching low order Taylor expansions [128] and interpreting a 3-D image as an instance of a surface in 4-D space [36].

2.3 Rigid-body Motion Correction in FMRI

At present, most intra-modal 3-dimensional FMRI motion correction is performed assuming a rigid 6 degree-of-freedom (DOF) transformation between displaced images. This can be regarded as a reasonable approximation in view of the relatively small displacements and rotations assumed to be present in functional data sets. In addition, any physical motion of the brain must be bounded by the skull which is a rigid object and within which the brain does not deform appreciably. Even in the case of clinical patients, physical constraints due to the dimensions of the scanner mean that motions of a few centimetres at worst can be expected in the data. It is for this reason that, presently, linear registration models dominate work in this field.

2.3.1 Existing Voxel Intensity-based Methods

To date the two most widely used motion correction tools are Automated Image Registration [146, 145] (AIR) and the realignment algorithms included in the Statistical Parametric Mapping (SPM) package [44, 93].

AIR was originally developed for the realignment of Positron Emission Tomographic (PET) images[4] but has since been adapted to inter-modal registration problems (such as PET to MRI), inter-subject realignment including atlasing

[4] A technique for measuring the concentrations of positron-emitting radioisotopes within the tissue of living subjects.

and more specifically rigid-body registration of EPI data [81]. The cost functions offered for the purpose of rigid-body registration are the ratio of intensity uniformity (RIU), least squares (LS) and least squares with global rescaling. In addition, AIR supports a number of rigid-body and non-rigid models for the purpose of inter-subject registration.

SPM is a complete pre-processing and analysis environment for functional neuroimaging data, most commonly FMRI. The realignment procedure included in the software performs rigid-body realignment by estimating the 6 motion parameters through the minimisation of a LS cost function between each image of the sequence and a reference image, either the first image in the series or an average over all the images being co-registered. The minimum is searched for iteratively using a Gauss-Newton algorithm [120].

2.3.2 Contour-based Matching

An alternative both to anatomical landmark- and voxel intensity-based approaches proposed by Biswal *et al.* is to use contours extracted from the image. In [13] contour images are generated using a Laplacian of a Gaussian filter and edges are detected at the zero-crossings using a 3x3 window so that only closed-loop lines are detected. The cross-correlation coefficient, calculated on the contour images made from the two images being matched, is maximised to ensure a close match.

The authors assume motion between adjacent EPI images will be sub pixel so the cross-correlation is calculated in the Fourier domain to speed up implementation and the shift theorem is used to enable sub pixel registration. This also eliminates the problem of blurring contours present in interpolation schemes.

Work showed that large initial motion estimate offsets lead to local minima traps, presumably due to incorrectly matching to the edges of phantom images.

It was also established that the contour-based cross-correlation had a lower standard deviation and a higher absolute value than its intensity-derived equivalent, indicating a better match for the contour-based method on the test data used.

Biswal *et al.* note that local changes in signal activity mean that intensity-based methods will not be able to align images even when the registration criteria are satisfied, and distortions would occur for activation and resting-state data. They suggest that a contour-based registration scheme sensitive to brain features would be useful for inter-modality registrations/corrections. These issues are addressed later in section 4.1.

The question remains, however, as to whether a consistent set of contours can be robustly extracted from a low-resolution EPI image and how to ensure that edges corresponding to activation artefacts, which will be present only in those images acquired when the patient is exposed to a stimulus, are not detected.

2.3.3 Raw K-space Corrections

It is also possible to correct for motion artefacts by suitably transforming the k-space data before they are converted into an intensity image. One specifically k-space strategy is to use Navigator Echoes which are able to measure motion in one direction (1D) [3]. In [47], Orbital Navigator Echoes (ONE), which have circular k-space trajectories, are used to measure in-plane rotation and multi-axis (2-D in-plane) translational global motion.

Displacements are indicated by a phase difference between the current and reference echo while shifts in the magnitude profile of the echo with respect to the reference echo reflect rotational patient motion. The signal of an ONE for an object at a perturbed position after being rotated through α and displaced by (x_0, y_0) is

$$S'(k_p, \theta) = S(k_p, \theta - \alpha)e^{ik_p(x_0 \cos\theta + y_0 \sin\theta)} \tag{2.2}$$

where k_p and θ are polar coordinates (radius and azimuthal angle) for the trajectory in k-space. Such as displacement is illustrated in figure 2.4.

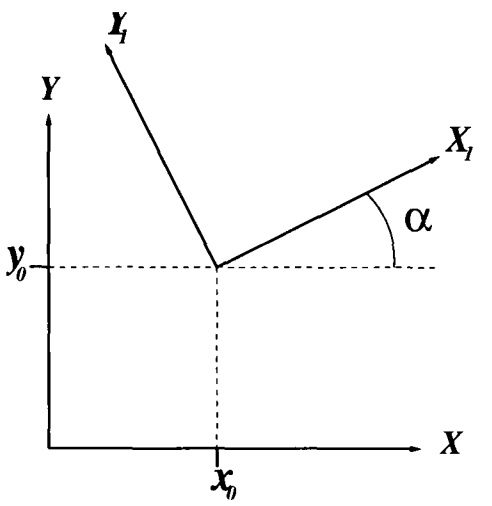


Figure 2.4: Coordinates of a displaced Orbital Navigator Echo (ONE) in k-space.

Detection of rotational motion is insensitive to any concurrent translations and therefore global view-to-view motion of an object (rotation and 2-D translation) can be determined by analysing the raw data of a single circular ONE. Corrections are applied using what the authors refer to as 'adaptive motion correction techniques' to suppress artefacts and blurring in both phantom and patient MR studies. This consists of translational re-sampling implemented by multiplying raw data with a phase factor determined by the corresponding displacement and k_y value ($k_y = k_p \sin\theta$). Attempts were made to correct rotational motion adaptively but this proved more complicated because of the angular mismatch between k-space points actually measured and those presumed measured.

Phantom studies show an almost perfect fit between navigator echo-derived rotational motion measurement and actual rotation (when these two sets of values

were plotted against each other, a slope of 1.02 ± 0.02 was given). The smallest angle measurement available was $0.7°$ corresponding to half a pixel of the navigator echo while the smallest displacement measured was half a pixel. However gradient errors could distort the orbital trajectory or cause the centre of the orbit to be displaced slightly from the origin of k-space. While these corrections are appreciable, sub-pixel motion is still significant so this method cannot be relied on exclusively for motion correction.

2.3.4 Real-time Corrections

In order to speed up the computation times of motion calculations, Cox and Jesmanowicz [28] compute a shear factorisation of the rotation/shift 3-D matrix and then apply these using two 1D fast Fourier Transforms (FFT). Making the assumption that FMRI requires only small rotational ($1° - 2°$) and translational (1–2 voxel dimensions) corrections, gradient descent is used to minimise a least squares (LS) objective cost function. In order to register the base image $J(\mathbf{x})$ to a target image $I(\mathbf{x})$, the LS incorporates a weighting functions as follows

$$E(\mathbf{a}) = \sum w(\mathbf{x}) \left[J(\mathbf{T}[\mathbf{a}]\,\mathbf{x}) - I(\mathbf{x}) \right]^2 \tag{2.3}$$

where $\mathbf{a} \in \mathbf{R}^6$ are the motion parameters, $\mathbf{T}[\mathbf{a}]$ is the spatial transformation corresponding to $\mathbf{a}$ and $w(\mathbf{x})$ is a smoothed version of the base image $J(\mathbf{x})$. The resulting motion estimates were within $0.05°$ and $0.04mm$ of AIR 3.08 [148] but ran 4.3 times faster. This agreement is unsurprising as both methods use a gradient-descent LS scheme.

The authors noted that faster CPUs could perform real-time registration with Fourier interpolation on complex-valued images, preserving the phase information

and maintaining full bandwidth of MRI data without aliasing. They also comment on the value of real-time feedback of motion estimates during acquisition so that overly-large estimates can prompt re-acquisition of data and to give patients some indication of how much they are moving.

Chapter 4 contains a more exhaustive survey of such real-time prospective schemes which seek to correct motion during the experiment. It is unlikely, however, particularly in the case of clinical patients, that feedback of motion levels would guarantee any reduction in subsequent motion. A further drawback of such schemes is that they require additional hardware in order to facilitate real-time motion correction. This is both expensive and useless in the correction of existing data-sets.

2.3.5 Slice-to-volume Registration

The rigid-body registration approach based on stacked slices (that is, assuming that all the slices acquired in each volume are aligned despite being acquired sequentially) may hinder statistical accuracy by introducing inaccurate assumptions of no motion between slices for multi-slice FMRI data. The slice-stack approach is inaccurate since each slice is excited sequentially.

Single-shot EPI (as opposed to multi-shot, also called interleaved or segmented EPI [23]) is the typical EPI sequence used in FMRI. The 'single shot' refers to the fact that each slice is reconstructed from a single RF excitation. It is possible to acquire several slices with single-shot (one shot for each slice). The FMRI data considered in this work are nearly always single-shot multi-slice EPI.

Other types of EPI include 3-D EPI and EVI (Echo Volumetric Imaging). In 3-D EPI, a slab (block of slices) is excited and the separation in the slice direction is determined by phase encoding. As many slice-phase-encode steps as

the desired number of slices are acquired, and when the data has been captured, a 3-D Fourier Transform is applied to obtain the image. The main disadvantages of this technique are wraparound in the slice direction and usually a lower SNR than a multi-slice acquisition. This is not always the case, but given the kind of slice thickness, coverage and TR common in FMRI, it has not been a feasible option to date.

EVI is the full extension of EPI to 3-D which means that the data are acquired in a single shot. It is extremely difficult to carry out in practice[5] and results in a very low-resolution image (typically 32x32x16 voxels) with large levels of distortion.

Because of this practical restriction to single-shot EPI and the possibility of inter-slice motion, Kim *et al.* adopt a map-slice-to-volume approach [86]. Individual slices acquired by a multi-slice single-shot EPI are repositioned into an anatomical reference volume space and reconstructed into volume data. Mutual Information is used as a cost function in the automatic registration necessary for correction of slices in EPI data sets containing thousands of slice images.

The Nelder-Mead downhill simplex optimisation algorithm [120] is used to drive the automatic registration. A mutual information metric is used to register slices and trilinear interpolation (which in turn uses the original data set for each iteration to avoid rounding errors) is used to calculate the new, corrected sub-voxel values in the corrected image. The 2-D grey value scatter plot of $p(x, y)$ (the joint histogram) is expected to give minimal dispersion when the two images are registered [87]. Although the routine is able to cope with 2-D/3-D affine and warping models, only a 6 parameter (rigid-body) model was implemented.

SPM approximately models slice-stack registration using least squares align-

[5]Personal communication with Dr S Clare, FMRIB.

ment which, Kim *et al.* claim, cannot offer as accurate a registration solution. Variability in subject motion is not assumed to follow any pattern, nor to be in-plane. It follows that voxels in registered slice-fit-into-volumes in each activation cycle may not match in sample counts as each slice is treated independently. This means that there may be areas in the realigned volumes which contain no data or conflicting data taken from different slices.

Results from tests carried out in [86] on simulated data show that the sub-voxel accuracy of the registration parameters is in the range of $-0.10° \pm 0.07°$ and $-0.07 \pm 0.03mm$ (expressed as the mean $\pm$ standard error of means) in maximum rotation and translation respectively. Unfortunately these results are unlikely to be representative of the algorithm's performance on real data which is likely to contain artefacts such as high field B_0 distortions. In general, 2-D to 3-D registration at these resolutions will not be robust, further decreasing the accuracy of such a method.

2.4 Validation of Motion Correction

In order to carry out a quantitative comparison of the range of realignment methods which have been described above, a set of measures which characterise the level of motion correction must be devised. While other applications in image analysis, such as segmentation, can be verified by recourse to expert opinion, this approach is not feasible for FMRI motion correction. First, as discussed in Chapter 1, the amount of data processed makes it impractical for an expert to check laboriously each successive pair of images. Second, the low levels of motion and poor resolution combine to make subject movement very difficult to assess accurately, even by an expert.

The most common approach to registration validation is to devise a set of data where a known motion is deliberately introduced enabling a numerical comparison with the motion parameters calculated by realignment. Such approaches are described in more detail in the following sections.

2.4.1 Established Measures of Registration Accuracy

Concerning a computed registration, there is no gold standard for accuracy regarding clinical practice. There is a widespread search for measures that somehow quantify registration accuracy. One cannot, with absolute certainty, quantify local registration errors but it may eventually be possible to say that it is unlikely for the error to exceed a certain bound. In many instances, detailed accuracy studies are only just starting. This has been delayed by the improper use of the terms 'accuracy', 'precision reliability' and 'robustness' in many studies.

Robustness or stability refers to the basic requirement that small variations in the input should result in small variations in the output. Reliability is the requirement that the algorithm should behave as expected.

Precision is defined as the typical systematic error that can be obtained when the registration algorithm is supplied with inputs. Precision measures can be obtained concerning the entire registration system, or applying to specific components (that is, patient, acquisition or paradigm).

Accuracy is a more direct measure, referring to the actual, 'true' error, and applies to specific registration instances. It can be divided into qualitative and quantitative forms. The latter is unavailable in clinical practice. Evaluation of registration concerning accuracy and precision is carried out at several levels: software phantoms, where the images simulate the clinical acquisition, provide us with ground truths and control over image degrading factors; physical phantoms

which make use of true image acquisitions, although the ground truth now has to be estimated; cadaver studies; studies using real patient data with a database of images containing generic as well as exceptional cases; and finally at the clinical level, where validation should be turned over to the clinicians involved.

There is a definite shift from extrinsic to intrinsic methods in research, with surface-based methods being overtaken by 'full image content' intensity-based methods. The latter are beginning to set the standard for registration accuracy: rapid increases in available computing power and the growing realisation that it is necessary to retain all the information available for an accurate registration have led to the growing dominance of intensity-based methods.

2.4.2 Comparison Studies

West *et al.* [143] have carried out studies of accuracy based on a gold standard of registration. CT and PET are separately registered to MR using the rigid-body transformation that minimised the mean square distance between corresponding fiducials in the two images. *Fiducial registration error* (FRE) is defined as the root mean square (RMS) distance between corresponding fiducials after registration and transformation. Another established metric is the *target registration error* (TRE) which is the error introduced in identifying a target point because of the inherent errors in locating the point used to calculate the registration transform.

The original data included fiducial markers which were then 'air-brushed' out and posted to an ftp site for analysis by a number of groups who were responsible for the development of different registration methods.

The results of the participating laboratories' methods were compared with this calculated gold standard set. Participants were asked to provided a set of 8

points for reasons of accuracy and error checking although only 3 are necessary to specify uniquely a rigid transformation, the limiting criterion of the study.

Methods submitted included: multi-resolution Powell algorithm minimising the Euclidean distance between two surfaces [96]; fully automatic Powell algorithm searching for the maximal mutual information of the two intensities [24]; fully automated multi-resolution grey-value correlation technique (only for CT to MR) [138]; surface-based method using a triangular mesh for the first image and a set of 2-D points representing the surface in the second which is subsequently subjected to a cost function related to the sum-square perpendicular distance between the two surfaces [61]; fully-automated multi-resolution voxel similarity measure based on the mutual information of the joint probability distribution [66, 65, 132]; fully automated matching 'edgeness' volumes [99, 101, 100]; segmentation followed by the construction of a potential energy field in one volume and the computation of the resulting global force and torque in the second [103, 104, 105]; obliquely transforming the image slices into coincidence followed by a first- or second-order polynomial (warping) transformation which seeks to minimise the RMS distance between point sets [83]; segmentation of contours followed by a Powell minimisation of the mean square distance of the points from the surface [117]; and a multi-sampling density method where a Newton-Raphson method is used to find the transformation parameters that minimise a weighted average of the standard deviations of PET voxel intensities corresponding to each MR partition [148].

Results suggested that the accuracy of some of the retrospective techniques approached the accuracy of bone-implanted fiducial marker methods. It is not possible to draw statistically meaningful conclusions regarding the superiority of any individual technique for a given task but this was always beyond the scope

of the project. The authors also note that visual inspection is necessary to guard against large errors, for example, allowing for new initial conditions or algorithm parameters.

These validation metrics cannot easily be applied to functional data. Most fiducial markers will not show up in a functional scan and cannot be located accurately. Furthermore, there are no reliable functional landmarks as such because the intensity features on a functional map fluctuate according to the presence or absence of the experimental stimulus.

This leaves ground truth to be established either through the use of phantom studies or by examining the performance of correction schemes on artificial data sets where the true motion is known exactly. Alternatively, some measure which quantitatively describes the degree of motion in a functional data set must be designed and then compared with the same measure taken after motion correction.

Assuming that the accuracy of such schemes has been verified on synthetic data, the ultimate validation is still that the results are not only clinically useful but are trusted. Specifically FMRI data must indicate (in)activity in regions which can be correlated with other independent information, whether it be data from another modality such as PET, Magnetoencephalography[6] (MEG) or Electroencephalography[7] (EEG); or simply from clinical observation.

2.5 Algorithm Design

The first aim of the work described in Chapter 3 is to design and implement a rigid-body motion correction scheme which is specific to FMRI data. The nature

[6]Which measures the intercellular currents of neurons in the brain.

[7]Where electrical signals produced by the brain are measured via electrodes attached to the scalp.

of the algorithm is determined by three basic considerations: a retrospective correction; modality-specific; and performance. These are justified briefly below.

2.5.1 Retrospective Correction

In choosing to develop a retrospective motion correction scheme, rather than a prospective approach such as the one described in section 2.3.4, it is possible to apply the method to any FMRI data already existing without the need to re-scan subjects with specialist acquisition sequences or additional expensive hardware.

A further advantage, intrinsic to a retrospective scheme where the input is the Fourier transformed image, is that direct feedback is possible between observations regarding the spatial manifestation of motion artefacts in the data and image-based registration methods which seek to correct for these confounds.

2.5.2 Modality Tuning

FMRI is an intrinsically low resolution modality, which does not support the use of landmark-based schemes which might otherwise aid in registration. The datasets presented for correction contain a large number of scans where the majority of voxels represent areas with low, or even non-existent, response to experimental stimuli.

By biasing a cost function towards these steady-state areas, for example, there will potentially be many areas within the images being co-registered which exhibit broadly similar intensity patterns across all scans. This suggests an intensity-based approach in combination with a similarity cost function as the intensity and contrast of the images will be consistent within tissue types.

In section 2.3, it was noted that a rigid-body parameter model should be

sufficient to describe the motion in FMRI data. There are several optimisation schemes such as gradient descent or Powell's method which can be applied to the parameter search.

It has been shown that accuracy studies may be performed on artificial data, with the consequent advantage of reflecting a motion which is known accurately. It is expected, however, that in addition to the varying activation patterns across the images, there will also be a number of additional confounds resulting from baseline physiological artefacts. These will necessitate a thorough evaluation of the scheme on real data leading to possibly qualitative conclusions.

2.5.3 Performance

The main goal of any motion correction scheme is to achieve accurate results and a robust performance over the full range of FMRI data to which it may be applied. Secondary to this is the need for rapid execution and minimal or no user interaction so that motion correction may be applied automatically to large data-sets to give consistent results.

Chapter 3

Rigid-body Motion Correction

The chapter begins with a review of the registration problem in the context of medical image analysis with particular emphasis on the FMRIB Linear Image Registration Technique (FLIRT) [74]. This forms the basis for the development of a novel approach to motion correction in FMRI, which will be called Motion Correction using FLIRT (MCFLIRT) [7, 73].

This new method has been applied both to real and artificial data-sets, demonstrating that it provides more accurate and robust results than other existing approaches to motion correction in FMRI. The results of this validation are detailed in the later sections of the chapter. In particular, direct comparisons are made between MCFLIRT and currently the two most widely used FMRI motion correction tools, AIR and SPM, both of which were described in section 2.3.1.

The chapter concludes by observing that although the corrections achieved with MCFLIRT are to date the most accurate available, residual motion artefacts remain in the data, so that a more sophisticated model of the acquisition and movement process is required for comprehensive motion correction. This is the subject of Chapters 4 and 5.

3.1 MCFLIRT Motion Correction

In functional brain imaging, a series of brain images is taken in quick succession, with a time lapse between each acquisition of usually a few seconds or less. Due to the small acquisition times, the images usually have poor resolution, as illustrated in the example shown in figure 1.2. Furthermore, as a consequence of the fact that the imaging parameters are chosen to highlight physiological changes (e.g. blood flow), the images often have poor contrast.

Normally, motion correction methods deal with the registration task by selecting a reference image from within the series and registering each image in turn to this fixed reference. As all images are of the same subject, using the same imaging parameters, the realignment be classified an intra-subject, intra-modal registration problem. Therefore a rigid-body transformation space and intra-modal cost function can be used. In addition, as the values in the corrected images are important for later statistical analysis, the choice of interpolation method for the transformation of the images is of particular importance [59, 58].

3.1.1 Problem Statement

The registration of two images I^r (nominally the reference image) and I^f (the target or 'floating' image) can be computed by constructing a cost function which quantifies the dissimilarity between two images and then searches for the transformation (T^*) which gives the minimum cost value. Mathematically this may be expressed as

$$T^* = \arg\min_{T \in S_T} C(I^r, T(I^f)) \tag{3.1}$$

where S_T is the space of allowable transformations, $C(I_1, I_2)$ is the cost function

and $T(I^f)$ represents the image I^f after it has been transformed by the transformation T.

The reference image may be of a different modality [13]; but a more common approach is to select one image from the time-series (usually the first, for example SPM [46]) and to register all the remaining images to this image.

3.1.2 Cost Functions

Intensity-based cost functions can be separated into two categories: those suitable for intra-modal problems and those suitable for inter-modal problems. In the former category the most commonly used cost functions are: *least squares* and *normalised correlation*. For the latter, and more difficult, category the most commonly used functions are: *mutual information*, *normalised mutual information*, *ratio of intensity uniformity* and *correlation ratio*. These functions are defined mathematically in figure 3.1.

A more detailed description of these cost measures is included in Appendix III. In the case of the entropy-based cost functions, it is common to adopt a windowing technique such as Parzen density estimation [33] to estimate the unknown probability densities.

Interpolation

The cost function requires the definition of an interpolation, i.e. a technique to calculate the intensity in the floating image I^f at points between the original grid (voxel) locations. The most common interpolation methods include: trilinear (also called linear or, in 2-D, bilinear), nearest neighbour, sinc (of various kernel sizes and with or without various windowing functions, for example, Blackman), spline and Fourier.

Cost Function	Definition	Minimum	Maximum
C^{LS}	$\sum(Y-X)^2$	0	∞
C^{NC}	$\frac{\sum(X.Y)}{\sqrt{\sum X^2}\sqrt{\sum Y^2}}$	-1	1
C^{RIU}	$\sum_k \frac{n_k}{N} \frac{\sqrt{\text{Var}(Y_k)}}{\mu(Y_k)}$	0	∞
C^{CR}	$\frac{1}{\text{Var}(Y)} \sum_k \frac{n_k}{N} \text{Var}(Y_k)$	0	1
C^{MI}	$H(X,Y) - H(X) - H(Y)$	$-\infty$	0
C^{NMI}	$\frac{H(X,Y)}{H(X)+H(Y)}$	0	0.5

Figure 3.1: Mathematical definitions of the most commonly used intensity-based cost functions. Note that the form of each metric, in particular the entropy-based measures, may differ from the similarity function formulation. The notation is: quantities X and Y denote images, each represented as a set of intensities; $\mu(A)$ is the mean of set A; $\text{Var}(A)$ is the variance of the set A; Y_k is the kth iso-set defined as the set of intensities in image Y at positions where the intensity in X is in the kth intensity bin; n_k is the number of elements in the set Y_k such that $N = \sum_k n_k$; $H(X,Y) = -\sum_{ij} p_{ij} \log p_{ij}$ is the standard entropy definition where p_{ij} represents the probability estimated using the (i,j) joint histogram bin, and similarly for the marginals, $H(X)$ and $H(Y)$. Note that the sums in the first two lines are taken over all corresponding voxels.

With the exception of nearest neighbour, all the listed interpolation methods are continuous. In the case of motion correction, however, it will be shown that the specific choice of interpolation method is crucial, since the intensities in the transformed images are needed by the later statistical analysis, outlined in section 1.3.

Regularisation of the Cost Functions

Research has shown that the cost functions exhibit small discontinuities as the transformation parameters are varied smoothly. This creates local minima 'traps' for the optimisation method. Since all interpolation methods are continuous (except for nearest neighbour, which is consequently seldom used) the discontinuities are due to the changing amount of overlap between the reference and the floating image.

It is possible to perform calculations strictly in the overlapping region but this leads to variation in the number of points counted in the overlapping region. This in turn means that both the numerator and denominator of certain cost functions, such as correlation ratio, will change discontinuously as the amount of overlap changes.

The discontinuities exist because images are finitely sampled and consist of discrete sets of voxels. In particular, the reference image defines a fixed set of voxel locations over which the cost function is calculated. Then, for a given transformation, the floating image intensities at these locations are calculated using interpolation. A reference image voxel location is counted only when it is valid: that is, within the overlapping region such that it maps to a location inside the field of view (FOV) of the floating image. When the edge of the FOV of the floating image crosses a reference voxel location, the location suddenly changes from being inside the overlapping region to outside, causing a discontinuous change in the number of valid locations, as shown in figure 3.2.

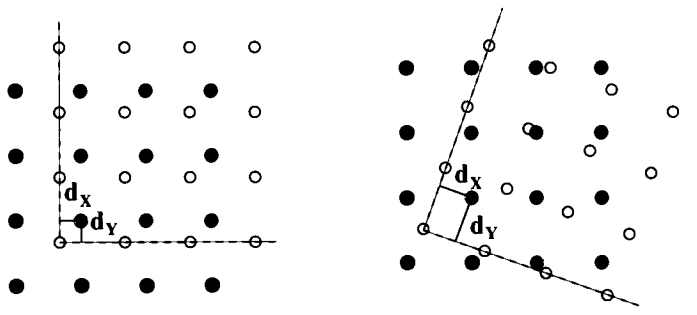


Figure 3.2: Example showing how the overlapping region of the FOV for the reference image (filled dots) and floating image (open dots) can change with transformation (translation and rotation on the left and right respectively). The shortest distances, d_X and d_Y, (shown as solid lines) to the edges of the overlapping region (shown by the dashed lines) are illustrated for a single reference voxel location. Those voxels which lie within one unit of the edge of the FOV, which are subjected to a weighting function, are shown in red while voxels completely inside the FOV are shown in blue.

Both FLIRT and MCFLIRT regularise[1] the cost function by removing such discontinuities via the introduction of geometric smoothing which reduces the weights contributed by locations that are near the edge of the overlapping region [73]. The weighting is chosen so that the contribution of such locations reduces monotonically until it reaches zero at the edge of the overlapping region. For simplicity and computational efficiency a linear weighting function is used, as shown in figure 3.3.

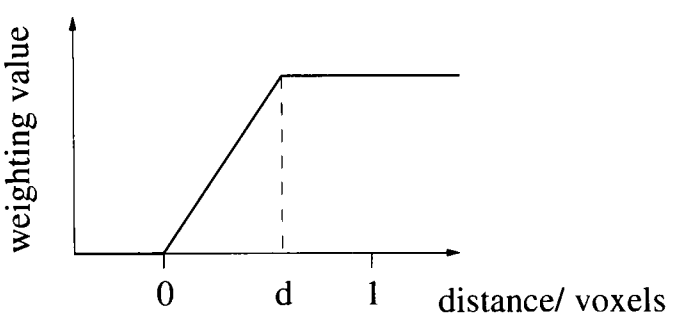


Figure 3.3: Voxels which fall within one voxel of the edge of the overlapping region of the FOV (extending from 0 to 1) are weighted according to a linear function which depends on the distance, d, from the boundary of the FOV.

For instance, consider a 2-D example of a reference location that maps to a point inside the overlapping region, where the distance from the nearest edges of the floating image FOV are d_X and d_Y units, as shown in figure 3.2. In each dimension, if this value is less than some threshold D, then the influence of that point is weighted by $w = d/D$. In higher dimensions, the product of the weights in each dimension is used. That is: $w(d_X, d_Y, d_Z) = w(d_X)w(d_Y)w(d_Z)$.

This weighting scheme can be applied simply and efficiently to any of the non-entropy-based cost functions (i.e. LS, NC, RIU and CR). It depends on a single parameter (the threshold distance D) which can be varied to increase the amount of regularisation. When $D = 0$ there is no regularisation, whereas increasing D

[1]Otherwise known as smoothing or apodisation.

creates increasingly smooth cost functions, although the cost function will be continuous for any non-zero value of D.

3.1.3 Optimisation

Assuming that the cost function which is being used does not lend itself easily to gradient calculations, the optimal method for function minimisation in 1-D is known as *golden section search*. It is possible to bracket a minimum of a non-singular function $f(x)$ when there is a triplet of points (a, b, c), where $a < b < c$, such that $f(b) < f(a)$ and $f(b) < f(c)$. In this case, the minimum of $f(x)$ lies in the interval (a, c). It can be shown that the optimal bracketing of this interval occurs when b lies a fractional distance 0.38197 from one end and 0.61803 from the other [120], otherwise known as the *golden section*.

Given a bracketing triplet (a, b, c) on $f(x)$, the algorithm proceeds by repeatedly replacing x, where x is chosen in proportion to the golden section to lie in the larger of the two intervals (a, b) or (b, c). If $f(x) < f(b)$ then x replaces the midpoint b and b becomes an endpoint. If $f(x) > f(b)$ then x replaces the respective endpoint and b remains the midpoint. An example of this bracketing is shown in figure 3.4.

The process of bracketing continues in this way until the distance between the two outer points is within a given tolerance.

In order to extend this method to multiple dimensions, it has been found in practice that it is sufficient to perform a series of 1-D golden section searches on each of the 6 rigid-body motion parameters in turn. Comparison was made with Powell's method which produces mutually conjugate search directions but this was found to offer no practical advantage in terms of accuracy. Testing also confirmed that the sequential search was robust to a range of starting positions

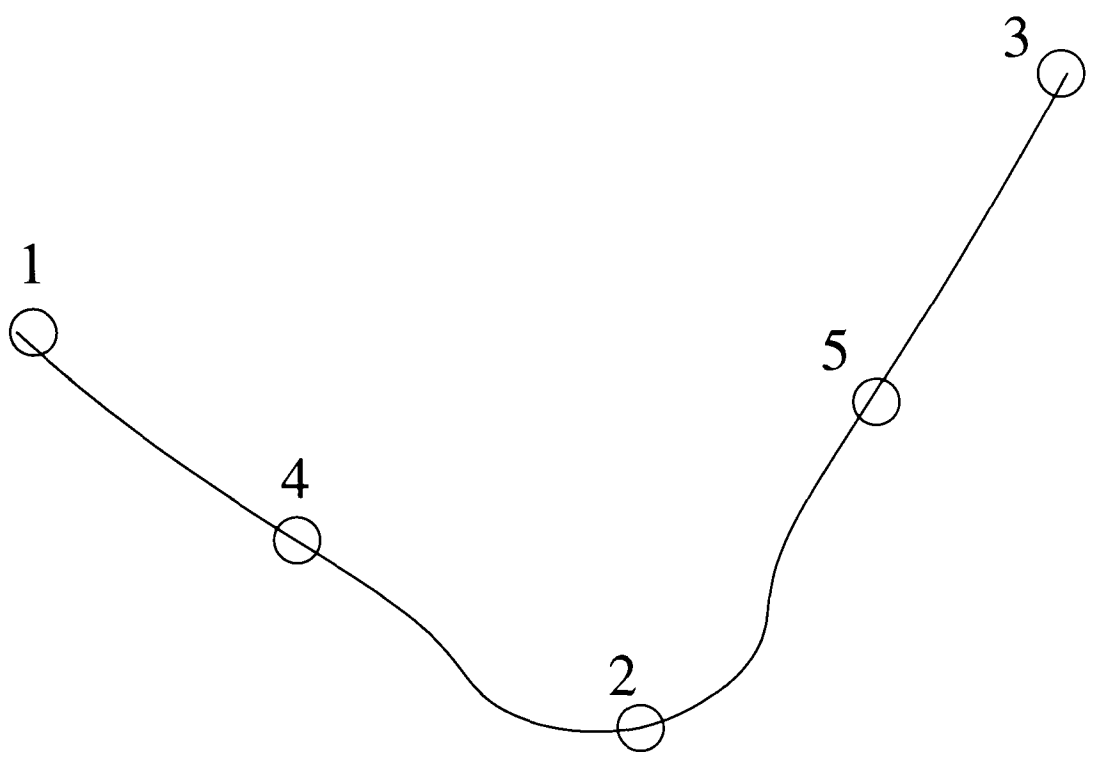


Figure 3.4: 1-D Golden Ratio search. Initial bracket $(1, 2, 3)$ becomes $(4, 2, 3)$, $(4, 2, 5)$ etc.

and bracketing sizes. The convention of first adjusting the rotational parameters was adopted as these will have a greater impact on the overall movement of the object than the translational parameters (for example, a rotation of 2 degrees centred in the middle of the brain will result in a translation of around 4 mm at the edge of the brain).

3.1.4 Local Initialisation

Making the realistic assumption that there is unlikely to be large motion from one image to the next, the parameters from the registration of one image can be used as an initial guess for the next image in the series. This is accomplished by assuming an initial identity transformation between the middle image in a time-series and the next adjacent image and then finding the optimal transformation by optimising the cost function. The resulting solution is then used as a starting point

for the next optimisation with the next image pair (see figure 3.5). This is done only at the lowest resolution, as all higher resolutions use the transformations found at the next lower resolution for the initial estimates.

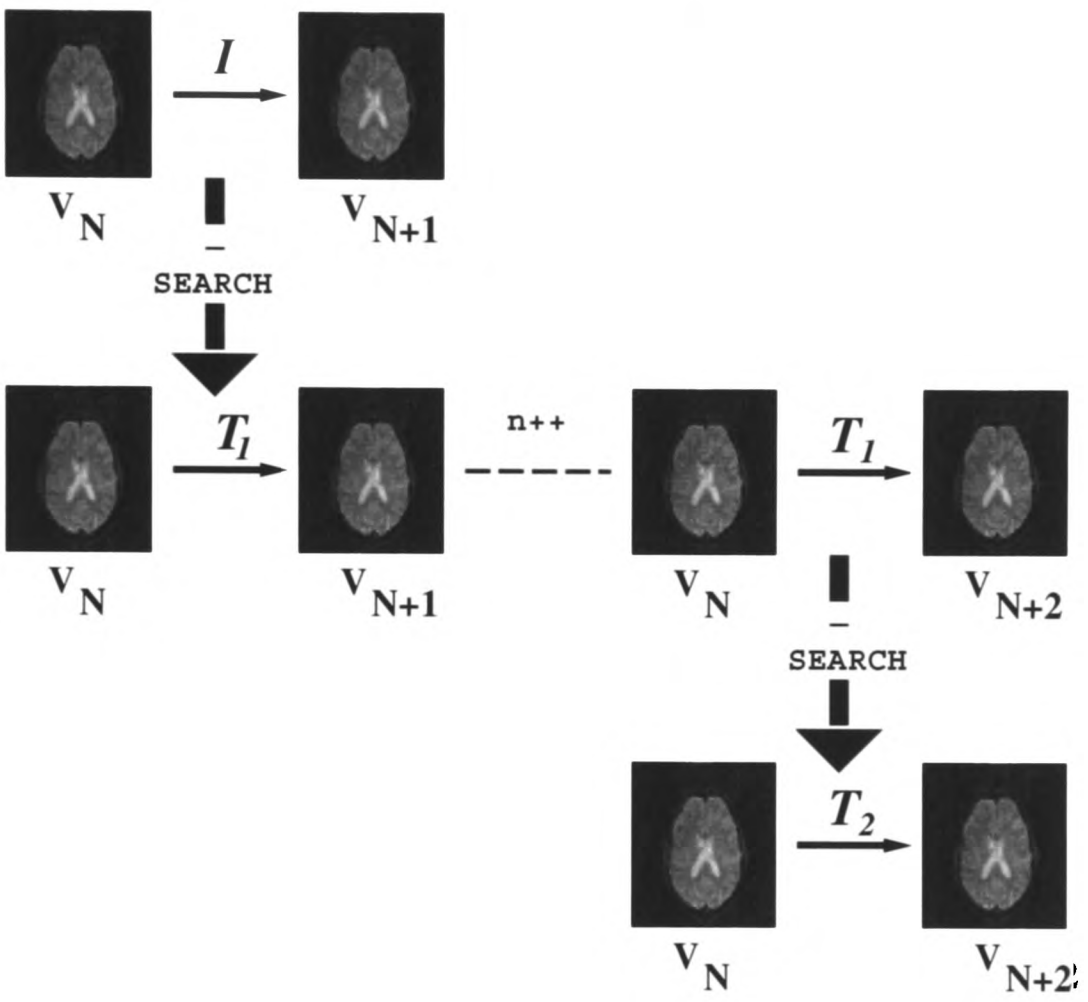


Figure 3.5: Schematic of the MCFLIRT rigid-body motion correction scheme. The middle image of the series is regarded as the reference image and each transformation to an adjacent image is used as the initial 'guess' for the transformation between that image and the one beyond.

3.1.5 End-slices

As the intensity values are of great interest after motion correction, attention must be paid not only to the estimation of the transformation but to its subsequent application. Interpolation probably has the largest impact on the quality of the transformed data, with sinc interpolation methods often being used, though there

is no consensus about which is the best method. Regardless, the loss of information outside the FOV, usually seen in the end-slices, can also be detrimental to the final statistical maps in these areas.

MCFLIRT has also been designed to handle the potentially problematic issue of end-slice interpolation. It is quite likely that under even small motion of the head, voxels at the top and bottom of the head can move either in or out of the field of view (see figure 3.6). Other schemes approach this by either assuming that all voxels outside the field of view and subsequently move in under registration are zero (AIR) or can be completely excluded from further calculations (SPM). This clearly has an impact on later analysis as valuable spatial information may be lost.

This situation can be avoided by 'padding' the end-slices. This is done by increasing the transverse (axial) extent of each volume by 2 slices and copying intensity values from the two end-slices onto their respective padding slices. The effect of this is that if data is to be interpolated from outside the FOV, it will take on 'sensible' values.

3.1.6 Method Summary

The final schedule in the form of pseudo-code is given as algorithm 1.

The algorithm describes a multi-scale approach to co-registration. Initially the data is sub-sampled to 8 millimetre isotropic voxels (line 6) to allow the gross intensity patterns in the test and reference images to drive the registration. The first registration pass is performed on line 11 using this scaling and implements the sequential registration described in section 3.1.4 with respect to the median volume in the data (line 3).

The actual cost function evaluation uses one of the metrics described in sec-

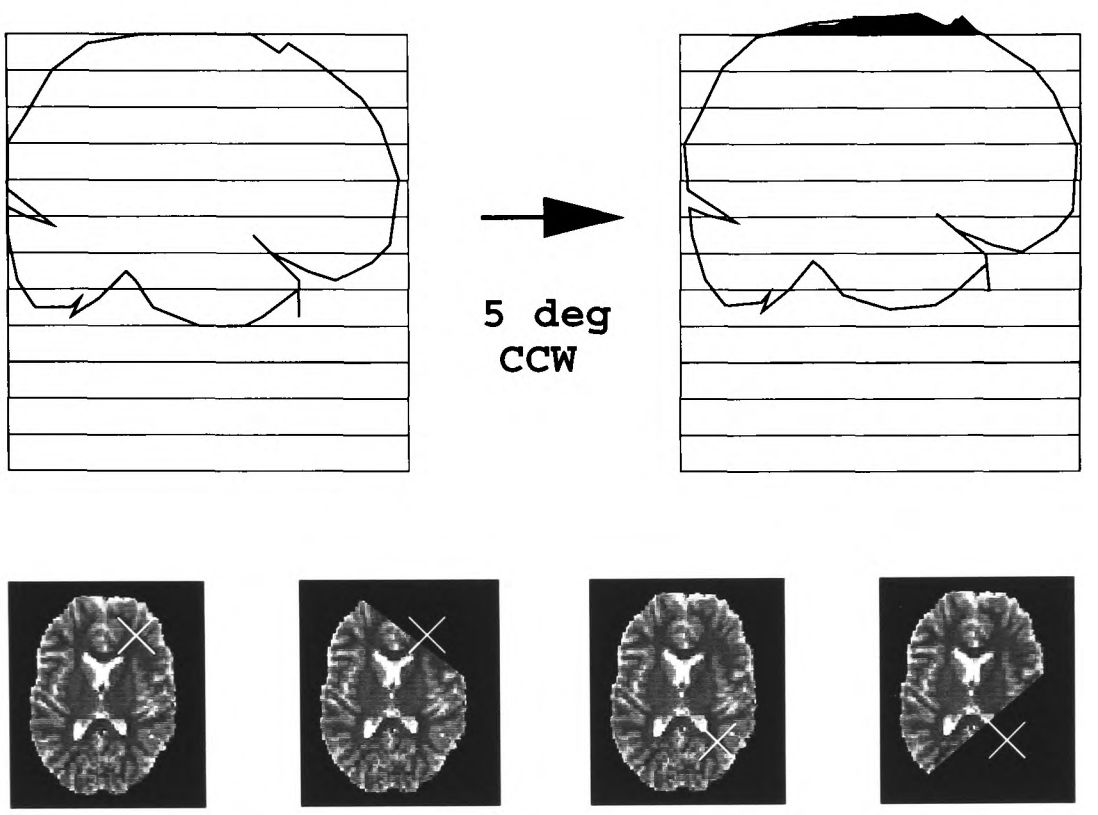


Figure 3.6: *Top row:* Demonstration of a significant section of brain moving out of the FOV (depicted as a stacked-slice grid) when the patient exhibits 5° pitch about the centre of the brain. *Bottom row:* Corresponding voxels in adjacent scans (shown as white crosses) can be lost if there is appreciable movement of brain in the end-slices. Diagram shows reference brain and corrected slice which has moved out of the FOV because of patient motion for two movements.

tion 3.1.2 and a comparison of the relative accuracies of the various costs available is presented below in section 3.2.1

Within the optimisation routine, each of the 6 rigid-body motion parameters is refined according to golden section search (section 3.1.3) until the tolerance specified on line 7 is reached.

Once the first pass is complete, there is the option to generate a new global reference image by computing the 3-D temporal mean of the corrected data (using trilinear interpolation) at this stage. This option is included on line 12 and the mean is generated as the volume:

Algorithm 1 MCFLIRT(time-series4d[ ] *ts*, Matrix[ ] *trans*)

```
1: time-series4d[ ] meanseries
2: n ←length(ts)
3: m ← n/2
4: vol3d refvol ← ts[m]
5: Matrix trans[m] ← Identity
6: scaling ← 8mm
7: tol ← large
8: Matrix init_trans ← Identity
9: for i ← 1 to n do
10:    if i ≠ m then
11:       trans[i] ← optimisation_trilinear(refvol, ts[i], scaling, init_trans, tol)
12: if mean_option = true then
13:    for i ← 1 to n do
14:       meanseries[i] ← apply_transformation(ts[i], trans[i])
15:    refvol ← temporal_average(meanseries)
16: scaling ← 4mm
17: for i ← 1 to n do
18:    if i ≠ m then
19:       trans[i] ← optimisation_trilinear(refvol, ts[i], scaling, trans[i], tol)
20: tol ← small
21: for i ← 1 to n do
22:    if i ≠ m then
23:       trans[i] ← optimisation_trilinear(refvol, ts[i], scaling, trans[i], tol)
24: if sinc_option = true then
25:    for i ← 1 to n do
26:       if i ≠ m then
27:          trans[i] ← optimisation_sinc(refvol, ts[i], scaling, trans[i], tol)
28: for i ← 1 to n do
29:    ts[i] ← apply_transformation(ts[i], trans[i])
```

$$\overline{\boldsymbol{V}} = \frac{1}{N} \sum_{\forall t} \boldsymbol{V_t}. \tag{3.2}$$

where N denotes the number of time-points in the data-set. This corresponds to line 15. If the mean option is selected then all subsequent optimisation runs (lines 19, 23 and 27) use this mean volume as the reference image. In the absence of the mean option, the optimisation runs conform to the description in sec-

tion 3.1.4 where the registration is carried out with respect to the middle image of the time-series (as before, selected on line 3).

For the second optimisation run (line 19), the images are evaluated at a higher resolution of 4 millimetre isotropic voxels (set on line 16). For the third pass using trilinear interpolation (line 23), this scaling is retained but the stopping tolerance associated with the parameter search is reduced (line 20).

If a fourth optimisation pass using sinc kernel interpolation has been specified, this is executed on line 27. Finally the transformation matrices, which have been progressively refined by these optimisation runs, are applied on line 29 using sinc interpolation.

The next section presents the results of testing the MCFLIRT implementation described in algorithm 1.

3.2 Results

This section presents several experiments which demonstrate the robustness and accuracy of MCFLIRT. After introducing some general test measures (which will be referenced throughout this evaluation), the implementation choices that have been made are described, as these are often critical in creating a stable method that performs well. Following this, registration experiments are presented which clearly demonstrate the improved robustness. The following sections demonstrate the improved accuracy brought to motion correction in FMRI.

3.2.1 Accuracy Assessment

This section details the comparative accuracy of the motion correction scheme (MCFLIRT) when tested against two of the most widely-used schemes: SPM and

AIR (see section 2.3.1).

Test Measure

If the images being used to establish registration accuracy have been synthetically generated, the motion parameters associated with this data-set will be known. It is therefore possible to compare the motion parameters estimated by the registration with those used to introduce motion artefacts into the test data.

In order to give a measure of the average error in position associated with a voxel in the image after correction (compared with the true measure of motion known for the synthetic data) a root-mean-square (RMS) average error metric is adopted [71]. Note that this metric will tend to favour those measures based on a least squares error minimisation, such as the cost function used in SPM.

The RMS measure is calculated to be the average difference in position for a voxel within the expected brain volume[2] between the generative matrices used to create the motion-corrupted data $\boldsymbol{T_1}$, and correction matrices produced by the motion correction scheme, $\boldsymbol{T_2}$. Given the low levels of intensity variability in FMRI data and hence the low occurrence of outliers, the RMS metric is deemed more suitable than alternative measures such as the L^∞ norm which characterise extreme error behaviour.

The RMS deviation in millimetres is calculated directly from the affine matrices, $\boldsymbol{T_1}$ and $\boldsymbol{T_2}$. That is:

$$d_{RMS} = \sqrt{\frac{1}{5}R^2 \,\mathrm{Trace}(\boldsymbol{A}^\top \boldsymbol{A}) + (\mathbf{t} + \boldsymbol{A}\mathbf{x_c})^\top (\mathbf{t} + \boldsymbol{A}\mathbf{x_c})}, \tag{3.3}$$

where R is a radius specifying the spherical volume of interest, $\mathbf{x_c}$ is the centre of the volume of interest and

[2]Taken to be a sphere.

$$M = T_2 T_1^{-1} - I \tag{3.4}$$

$$= \begin{bmatrix} A & \mathbf{t} \\ 000 & 0 \end{bmatrix}. \tag{3.5}$$

Later tests will also characterise the degree of correction by examining the residual error in the corrected data, described in more detail in section 3.2.1. This residual measure will also be used to demonstrate the scheme's effectiveness on real data where there is no absolute measure of the subject's movement.

Simulated Data

The artificial data enabling gold standard comparisons were generated as follows: a high-resolution EPI volume (with voxel dimensions of $2 \times 2 \times 2$ mm) was duplicated 180 times and each volume was transformed by an affine matrix corresponding to real motion estimates taken from a study, where the subject had been asked to move their head appreciably during the scan. This procedure was then repeated, using motion estimated from another data-set acquired under similar conditions.

Three further groups of images were generated using motion estimates from experiments where the subject had been asked to remain as still as possible. Within each of these five motion designs (two with appreciable motion and three without), three further groups of data were created to simulate audiovisual activation at 0%, 2.5% and 5% of the overall voxel intensities by modulating the intensity values according to an activation mask derived from real FMRI data.

Once the activation (if any) had been applied and the volumes transformed by

the corresponding parameters, the images were sub-sampled to 4 × x 4 × x 6 mm voxels. The use of a high-resolution template image which is then sub-sampled should minimise the effect of interpolation when applying such transformations to the data, as interpolation artefacts will be smoothed out by the sub-sampling.

Within MCFLIRT there are a number of stages which can be tuned to optimise the accuracy of the correction. The remainder of this section aims to find a robust set of parameters which give consistently accurate results on all data presented. It begins by examining the comparative accuracy of several cost functions which work within the proposed optimisation scheme. Later there is an examination of the impact delivered by the choice of interpolation scheme and registration schedule.

Cost Functions

The test results shown in figure 3.7 show the relative accuracy of the available cost functions within the FLIRT optimisation framework when applied to the problem of motion correction on our synthetic data.

Although there is no clear leader over all cost functions in terms of accuracy, it should be noted that the most accurate results are predominantly yielded by the Normalised Correlation and Correlation Ratio cost functions. This observation is reinforced by counting the instances where a particular cost function is most accurate is compared over all the available data-sets. This is summarised in figure 3.8.

It is interesting to note that previous work [42], which had demonstrated the superiority of entropy-based cost measures over alternatives in terms of motion correction without introducing further spurious activations in the data, has only compared Mutual Information metrics against least squares (SPM) and RIU

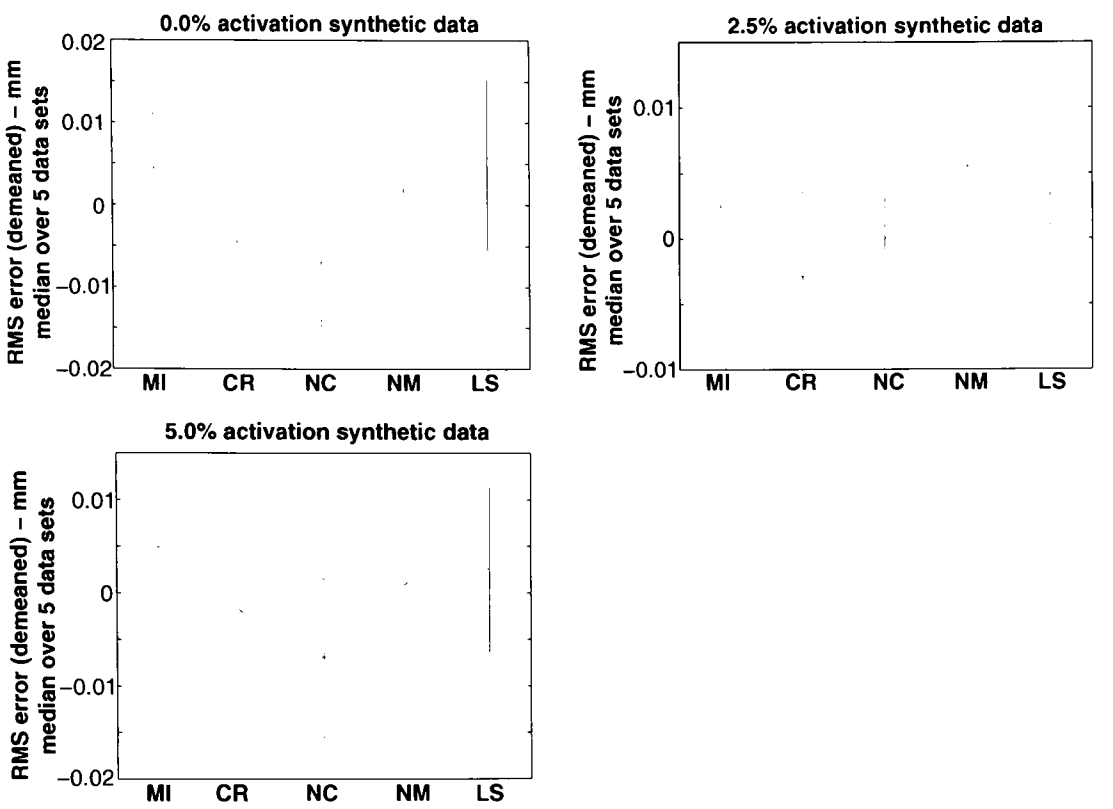


Figure 3.7: RMS median error results for the MCFLIRT Motion Correction scheme applied to synthetic data exhibiting known motion of one of five designs and audiovisual activation at increasing intensities. Error bars indicate minimum and maximum errors. Cost function notation corresponds to figure 3.1. Note that these RMS measures are relative: the original values exceeded the range of the display software and de-meaning was carried out across cost functions at each activation level. This accounts for the negative values presented in the figure.

Cost	# most accurate	# 2^{nd} most accurate
Normalised Correlation	8	5
Correlation Ratio	2	7
Mutual Information	2	1
Normalised Mutual Info.	0	2
Least Squares	3	0

Figure 3.8: Accuracy counts for the five cost functions offered by MCFLIRT applied to 15 data-sets described in section 3.2.1.

(AIR) measures.

Results using the RMS measure (figure 3.9) revealed that AIR 3.08 using Least Squares (which has been reported to give better results than the standard AIR RIU measure on FMRI data) and Windowed Sinc interpolation was almost

an order of magnitude worse than basic 3 stage trilinear MCFLIRT. Accordingly it was decided not to compare it further.

	Uncorrected	**AIR**	**SPM**	**MCFLIRT**
Sum of squared errors	936.5866	406.8876	1.6405	1.5171
Median error (mm)	2.3360	1.7570	0.1064	0.1102

Figure 3.9: RMS deviation values for synthetic null data across different motion correction schemes (AIR, SPM, MCFLIRT) compared with no correction.

It is salient to note, however, that AIR was originally developed for inter-modal registration and it is only because of its early prevalence in that field that researchers have sought to adapt it for intra-modal corrections, a task for which it had not been properly optimised.

Smoothing Option

The next stage of testing was to verify that these cost functions were in fact more accurate when regularised than un-regularised. The same RMS test measure was used as in the previous test and the data again exhibited one of the five motion designs at one of three activation levels. Results are given in figure 3.10.

Overall, and as expected, the smoothed cost functions outperform their un-smoothed versions.

Interpolation Scheme

To further improve the accuracy of the motion estimates, the next parameter to be varied was the choice of interpolation scheme for the motion estimation. In addition to the standard trilinear scheme, a windowed-sinc interpolation (using a Hanning window of width 3 voxels) was investigated.

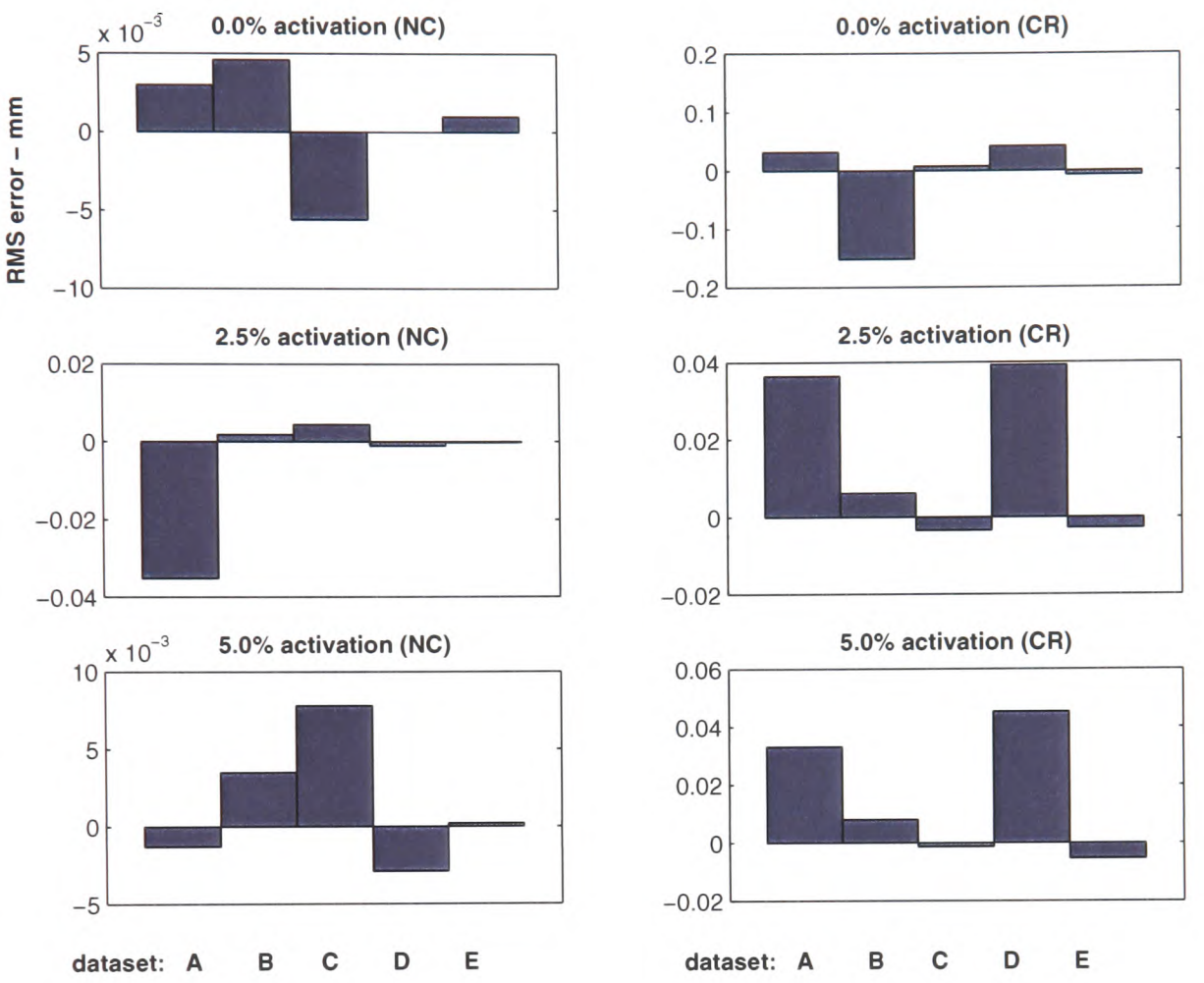


Figure 3.10: RMS median error results (improvement plotted as un-smoothed minus smoothed) for the MCFLIRT Motion Correction scheme applied to synthetic data exhibiting known motion at increasing intensities. A positive value indicates improved accuracy as a result of smoothing the cost function. Cost function notation corresponds to figure 3.1 and the results demonstrate the improvement in accuracy achieved by using the smoothed cost functions.

While considerably slower than trilinear interpolation, the sinc approach is able to refine motion estimates further after the initial trilinear stage has converged on a solution, thus providing greater accuracy. The results in figure 3.11 show the greater degree of accuracy achieved over using trilinear interpolation alone.

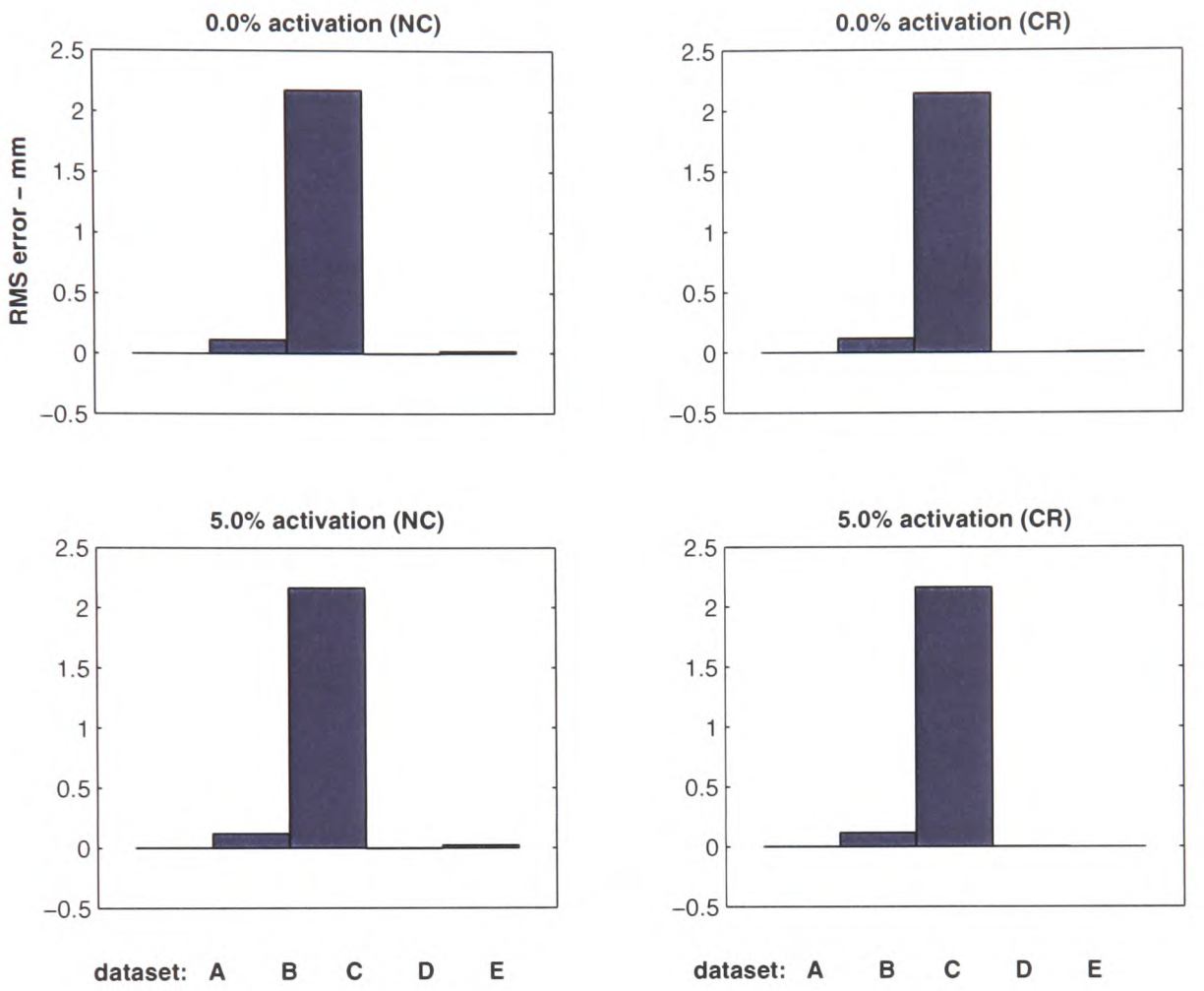


Figure 3.11: RMS error results (trilinear minus sinc final interpolation) for the MCFLIRT motion correction scheme applied to synthetic data exhibiting known motion of one of five designs and audiovisual activation at increasing intensities. A positive value indicates improved accuracy as a result of incorporating the final sinc interpolation stage. Cost function notation corresponds to figure 3.1 and demonstrates the improvement in accuracy achieved by using smoothed cost functions and additional sinc interpolation when compared to the basic trilinear scheme reported in figure 3.10.

Choice of Template Image

In an attempt to further increase the accuracy of the scheme and as a final parameter investigation, a method using a mean image template was implemented. This scheme generates a mean image for the series by averaging all the volumes over time after the first three stages of trilinear interpolation have been carried out. The intuition is that the target image for the registration of all volumes will be a more generalised image which exhibits less overall variation from each

volume in the series than the original target (middle) volume previously used. This new mean image is a robust target to which the original time-series is then registered, again using three stages of trilinear interpolation and an optional final sinc interpolation step.

Because registering includes a mean image, there are no longer gold standard values for the transformations found by the correction scheme. Usually a LS measure on the residual values would be suitable in this case but because SPM relies on the LS measure as its cost function, using the same metric as a test measure would bias the results.

Therefore, to quantify the accuracy of the correction, even without the RMS measure, a volume of the median absolute residual error (MARE) was created by, in each voxel, de-meaning the motion-corrected data and measuring the median value of the absolute residual values — effectively a measure characterising the level of inter-volume intensity variation (presumed to be due to subject motion) after retrospective motion correction had been applied.

While this can work only for activation-free data (so that in perfect alignment the variance is guaranteed to be at minimum) it can give a clear impression of the accuracy of the motion correction scheme. Because SPM rejects information outside a mask obtained from the data, the corrected median images were masked according to the corrected SPM data so that the measure reflected a consistent comparison across the schemes. The results shown in figure 3.12 correspond to the MARE values generated after running MCFLIRT and SPM on the null-activation data-set for both the low and severe motion designs.

From these results it can be concluded that for the low motion data-sets, MCFLIRT with the Correlation Ratio cost function provides the most consistently accurate results, while for high motion data-sets SPM99 provides slightly smaller

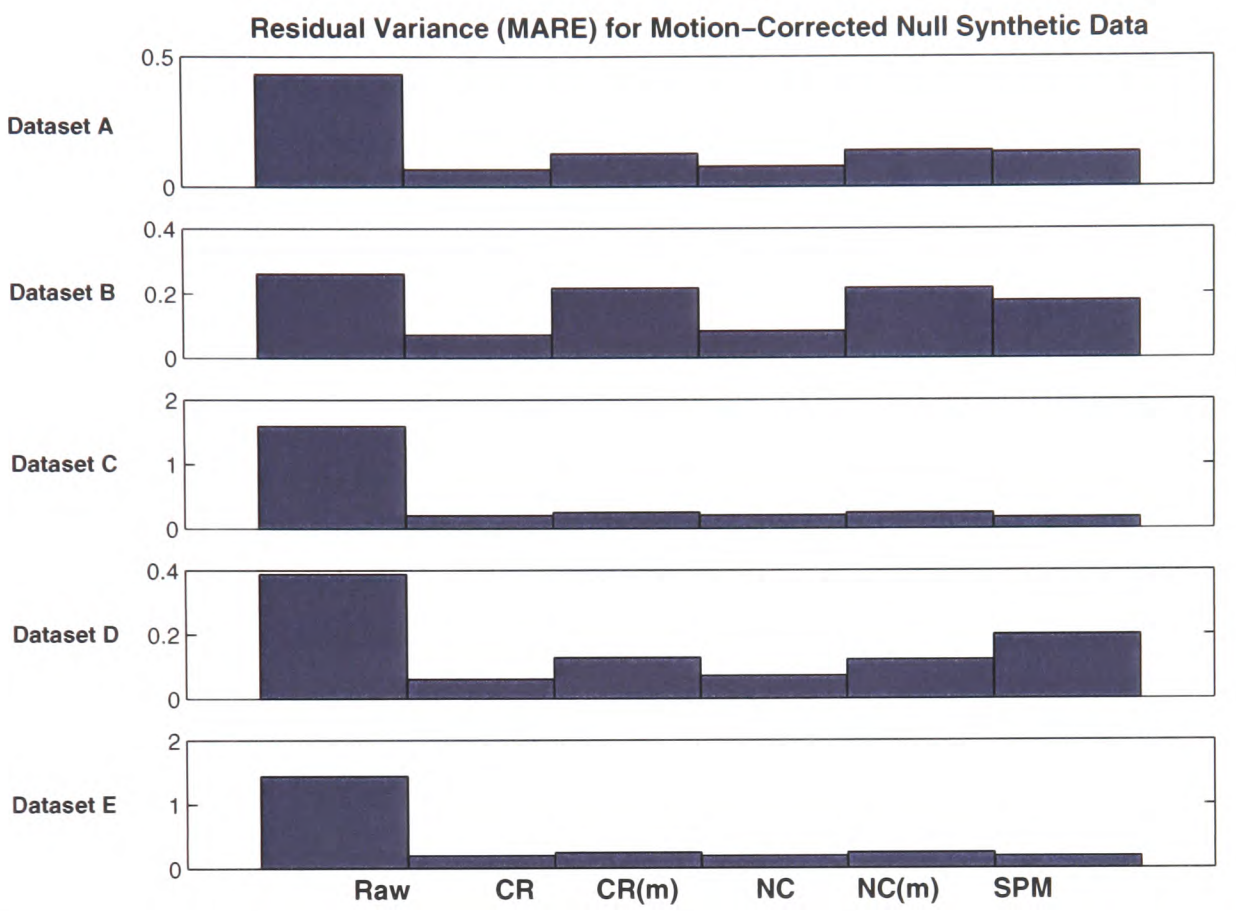


Figure 3.12: Median Absolute Residual Error (MARE) values for corrected data processed by different motion correction schemes: Uncorrected, MCFLIRT with CR, MCFLIRT with NC, MCFLIRT with CR and mean, MCFLIRT with NC and mean, SPM99 run at full quality, with sinc interpolation and interpolation error adjustment.

errors.

Slightly surprisingly it was found that the use of a mean image template gave no discernible improvement in accuracy; in fact, often the opposite. It can be concluded that for artificial data where the motion is purely rigid, there is no advantage to using a (possible blurred) average image over an image from the original data. In particular, the loss of edge information resulting from the averaging operation will be a significant factor in the reduced quality of the registration. One might expect the mean template scheme to yield greater accuracy where the data include non-linear artefacts and the choice of a reference image from such data-set is not so obvious.

3.2.2 Null Data Study

Having established the accuracy of MCFLIRT on artificial data, both MCFLIRT and SPM99 were applied to two real FMRI studies. In both instances, the subjects had been exposed to no stimulus (null data) but while in one instance the subject lay relatively motionless during the study, the second subject had attempted to move his head as much as possible within the confines of the scanner. This latter data-set exhibits motion of a magnitude close to that associated with a seizure. Again, results were masked according to the SPM data to give a fairer comparison. Results, presented in figure 3.13, show that for the extremely low motion data-set, both schemes lead to a minimal improvement with SPM marginally more accurate. In the case of the high motion data-set however, MCFLIRT offers a clear improvement over corrections using SPM. Qualitatively, the corrected images exhibit no visible motion.

3.2.3 Real Activation Study

With real data, it is difficult to make accurate measurements pertaining to the accuracy of a motion correction scheme. What has been done is to demonstrate that real data taken from a null study with high levels of motion exhibit lower residual error with respect to realignment after correction using MCFLIRT when compared with the results of SPM99 in figure 3.13.

With data exhibiting activation, examination of the time-series after correction using an animation tool reveals no obvious large affine movement although some motion artefacts remain. In particular, it is possible to show good localisation of activations which would not be possible without motion correction first being carried out. The thresholded statistical maps shown in figures 3.14 and 3.15

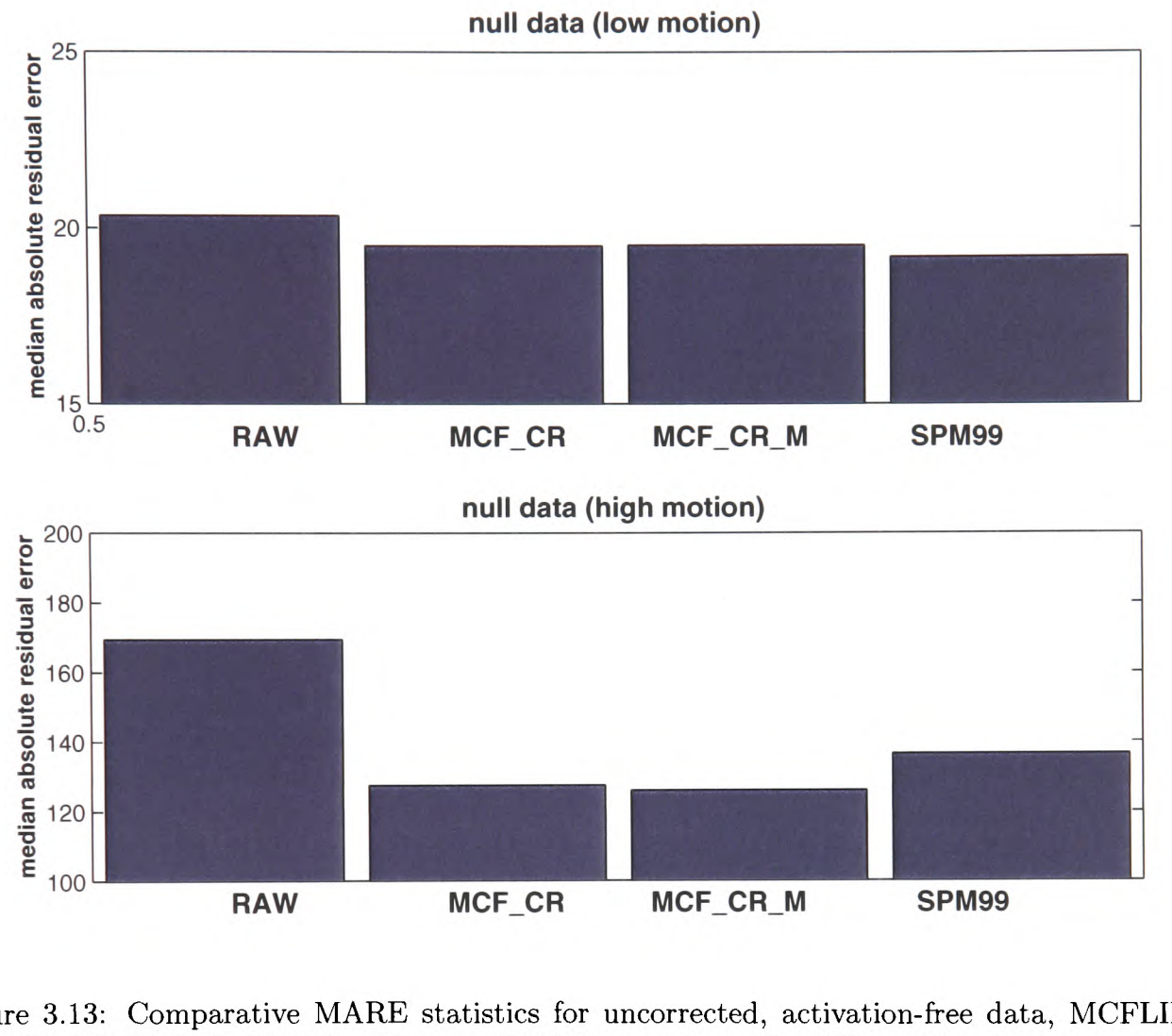


Figure 3.13: Comparative MARE statistics for uncorrected, activation-free data, MCFLIRT with CR, MCFLIRT with CR and mean, SPM99 correction using full quality, sinc interpolation and interpolation error adjustment.

correspond to a 180 volume audio-visual experiment[3].

Analysis was done using FEAT, FMRIB's Easy Analysis Tool using an Improved Linear Model [149]. In order to test the effectiveness of MCFLIRT on real data, the subject was asked to move his head during the experiment.

The uncorrected data lead to a thresholded z-map which contains a large number of false positives, given the nature of the experiment. While it can be argued that both data-sets are highly corrupted by this large motion (indeed even the corrected data-set still exhibits some visible movement, albeit on a signific-

[3]Where the visual stimulation was an 8Hz chequerboard (yellow and blue squares) with 30s OFF, 30s ON and the auditory was a recording of a mortgage discussion from Radio 4 with 45 s OFF, 45 s ON.

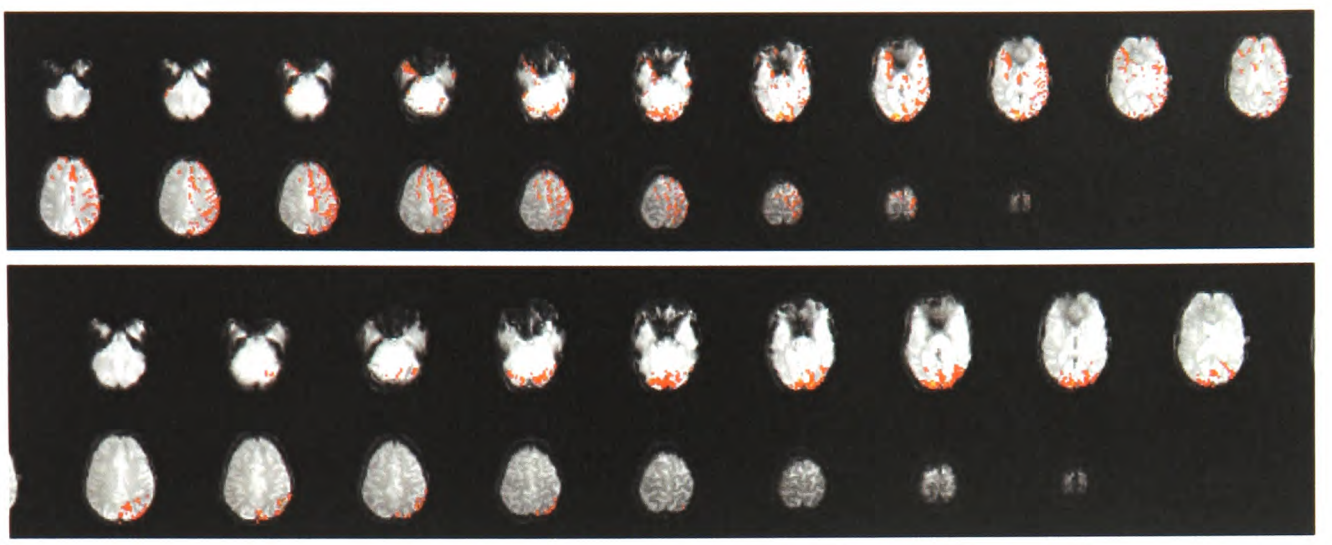


Figure 3.14: The resulting thresholded z-statistics from a FEAT analysis of a real audio-visual data-set (Auditory 45 s OFF, 45 s ON and visual 30s OFF, 30s ON) which has been motion corrected. Prior to motion correction, the series exhibited a large degree of motion because the subject had been requested to move as much as possible during the experiment. Maps are *(top)* visual paradigm z-statistics (uncorrected) and *(bottom)* visual z-statistics (MCFLIRT). It is expected that a visual response will manifest itself as an area of high activation (indicated in red) in the region of the visual cortex, at the back of the head. While the uncorrected data have incorrectly identified activation across the entire brain, the data processed by MCFLIRT display activation localisation as predicted.

antly smaller scale than the uncorrected data), the MCFLIRT-corrected visual activation is well localised and allows an otherwise corrupted set of experimental data to yield potentially useful results.

An ICA decomposition was also carried out on the corrected data and the results showed that of the set of new independent components, none exhibits the strong correlation to any of the six motion parameters present before motion correction. Some ICs show correlation to the motion parameters over a small number of time-points, which supports the idea that the ICA does not uniquely account for each motion parameter with a single component. This brief investigation demonstrates the use of ICA as a motion correction validation tool. It should also be noted that the ICs produced from the corrected data no longer exhibit 'rim-like' effects around the edge of the brain or ventricles which are associated

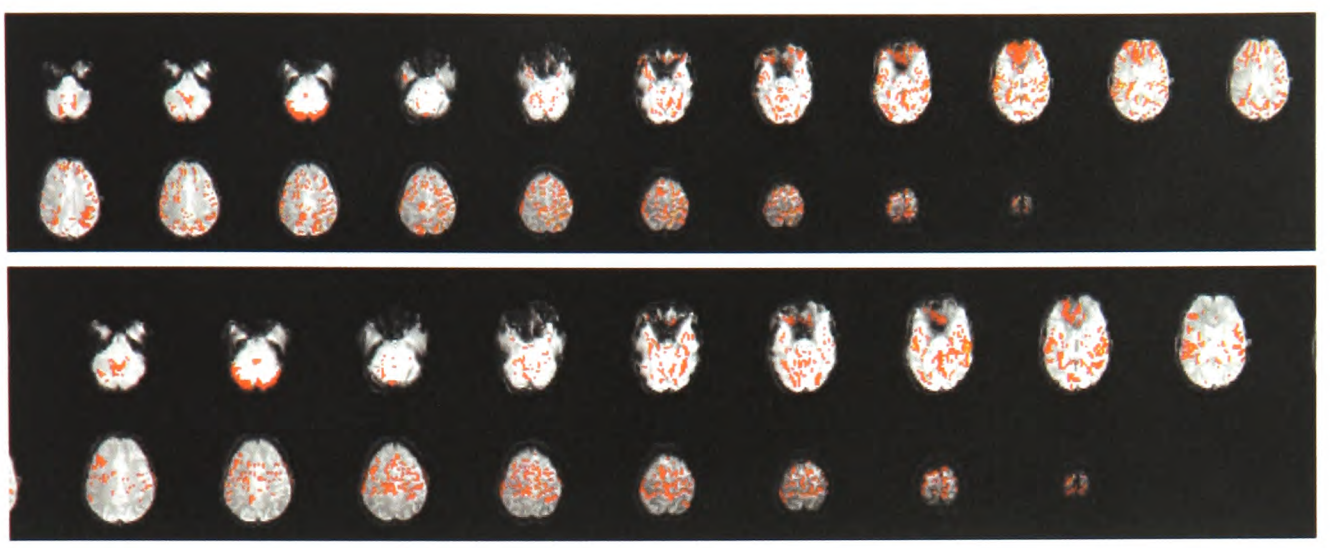


Figure 3.15: The resulting thresholded z-statistics from a FEAT analysis of the real audiovisual data shown in figure 3.14. Maps are *(top)* auditory paradigm z-statistics (uncorrected) and *(bottom)* auditory z-statistics (MCFLIRT). In the case of an auditory stimulus, the predicted response will be areas of high activation close to the ears themselves. This can be seen in the motion-corrected data with yellow peaks at either side of the brain while the uncorrected data show no clear spatial localisation of the activation response. The difference between the uncorrected data and corrected data is less distinct than in figure 3.14 but it must be borne in mind that the levels of data in this test experiment far exceed the levels expected from a normal, or even pathological, data-set.

both with slow- and fast-varying motion.

3.2.4 Discussion

In terms of accuracy, it has been shown that motion correction based on the FLIRT regularised cost functions, along with the sequential registration and sinc interpolation, consistently out-performs both SPM99 and AIR v3.08. It must be noted that all schemes are achieving sub-voxel accuracy as established using the RMS test measure. Such high levels of accuracy are necessary to ensure subsequent statistical analysis is valid.

Early tests using synthetic data have revealed that in cases where the motion is moderate (up to 2mm translation and 2° rotation), the sequential initialisation (see figure 3.5) scheme yields an improvement in the accuracy of motion estimates

compared with one where no sequential initialisation is performed. Conversely, in cases where the amplitude of motion parameters was known to be high, there was no inherent disadvantage in making the assumption of an underlying smooth motion trend across time-points.

It would be interesting to see how robustly the schemes perform over time-series of varying length. If at all significant, one might expect to see some impact on the MCFLIRT mean image registration scheme where a longer time-series might provide a more general and robust template image. At present there is no guaranteed advantage in using the mean template in addition to the standard correction schedule but one would expect it to play a more beneficial role in correcting extended time-series exhibiting moderate to low motion confounds.

3.2.5 Other Applications

MCFLIRT has also been modified to work with data which exhibit a restricted FOV, in particular single-slice data. In addition it has been reported to give good results for intra-modal registration of Diffusion Tensor Images[4] when the mean option is specified[5].

3.3 Residual Artefacts

Despite achieving sub-voxel accuracy in the corrected data, retrospective schemes still leave some residual motion and related artefacts uncorrected. By performing an Independent Component Analysis on the corrected data, it is possible to

[4]NMR images which measure the diffusion of water molecules within the brain. These images allow the development of techniques which can track physical neural pathways.

[5]Email from Erik-Jan Vlieger, Department of Medical Physics at the University of Amsterdam.

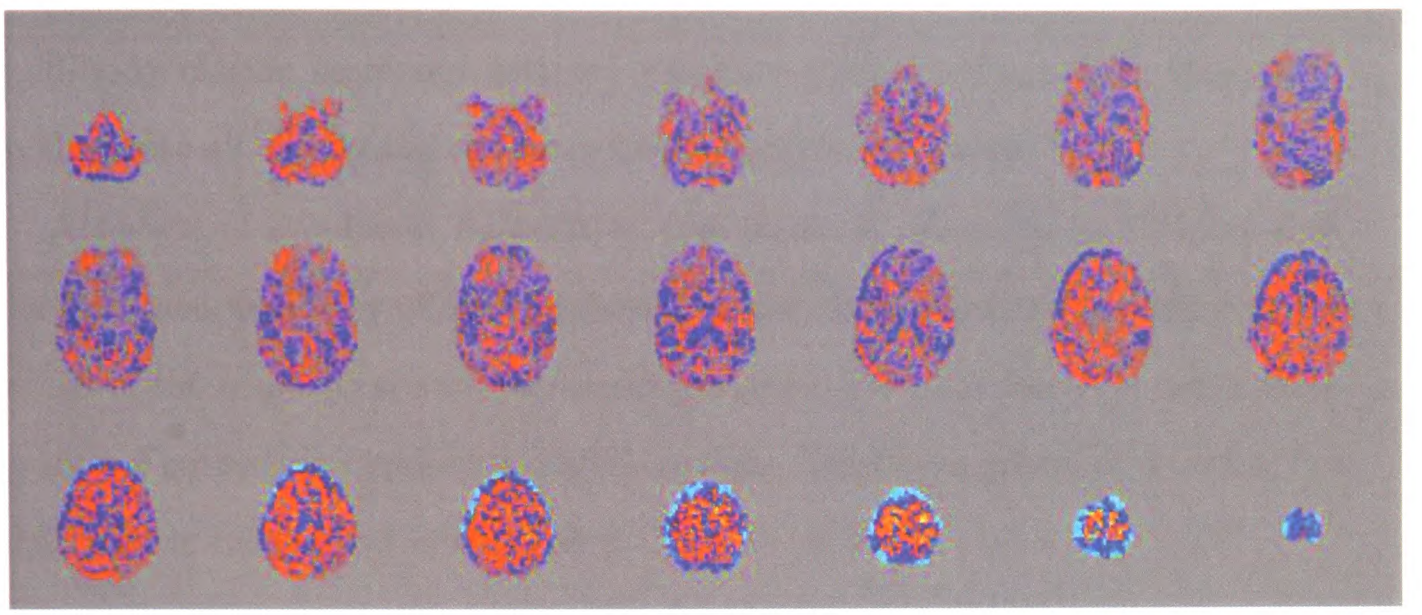


Figure 3.16: ICA spatial maps and time-courses showing residual motion artefacts in motion-corrected data. The strong edge effects in the top spatial map have led to this IC being labelled as a motion-related artefact.

identify a number of corrupting artefacts in the data which are remnants of the subject motion or in some cases introduced by the correction scheme itself.

These artefacts can be identified by examination of either the spatial or temporal components given by the ICA. An example is given in figure 3.16. It may be unrealistic to expect a rigid-body retrospective scheme to guarantee perfect realignment of images which may vary spatially due to activation-related signal changes or because of movement within an inhomogeneous magnetic field.

This suggests that despite the accuracy reported for MCFLIRT and other rigid-body motion correction schemes a more sophisticated approach is required to eliminate all the spatial artefacts resulting from movement.

A review of non-linear registration techniques is presented in Chapter 4 as the increased flexibility offered by these schemes, in combination with a detailed model of the acquisition and movement processes, are more likely to provide the quality of correction required in FMRI analysis. The development of an enhanced algorithm is the subject of Chapter 5.

3.4 MCFLIRT Summary

The work described in this chapter has led to the development of a rigid-body motion correction tool specifically designed for FMRI time-series analysis. The results which have been presented are able to demonstrate that this new algorithm improves the performance of existing methods for approaching the problem of subject motion in FMRI.

The motion correction package is available as part of the FMRIB Software Library (`www.fmrib.ox.ac.uk/fsl`) and typically runs in a few minutes on a FMRI data-set containing several hundred scans. At the time of writing the package is in use by more than 2000 researchers worldwide and published results have shown that it can be used to correct data-sets where SPM motion correction has previously failed [129].

Chapter 4

Non-rigid Motion Correction

While MCFLIRT has been shown to be fast, well-conditioned to FMRI and accurate, it was noted in section 3.3 that a number of artefacts remain in the images after motion correction. These have to be removed in order to ensure that accurate statistical inference can be performed on the data. As this is a problem common to the rigid-body schemes examined alongside MCFLIRT, alternative registration techniques need to be investigated.

There are a number of approaches which lend themselves to such improvements. They include non-rigid registration techniques, where matching is carried out using a more flexible data model, which it is hoped will be sufficient to represent the variability in the data introduced by non-linear artefacts. They also include prospective techniques, which employ a specially-designed acquisition sequence and/or custom hardware to allow artefacts to be removed while the object (in this case, the subject) is still in the scanner.

This chapter contains an examination of these techniques which, in the case of non-rigid retrospective methods, have already been applied to the registration of images from modalities other than FMRI. As with the earlier rigid-body registration review, it may prove in practice that the level of detail offered by FMRI

is insufficient to provide robust performance with some of these methods.

It is for this reason that prospective methods which can provide non-linear corrections are also evaluated. These methods have been developed specifically for MRI, in some cases FMRI, and while they necessitate additional levels of complexity at the acquisition stage, it may be that they are able to offer superior levels of motion correction to the more flexible intensity-based retrospective schemes mentioned above.

The chapter concludes by comparing the relative merits of these techniques with the aim of proposing a new model for FMRI data acquisition to be developed in Chapter 5.

4.1 Non-rigid Retrospective Registration

Internal motion, including significant variation in the movement of CSF boundaries throughout the brain during the cardiac cycle, implies that simply performing a rigid registration cannot be sufficient to eliminate completely artefacts caused by motion. This has prompted the investigation of a registration scheme designed for motion correction which employs a warping algorithm to determine a non-linear correspondence between images in a functional time-series.

Warps include global scaling, affine transformations and non-linear and local deformations. They can range from repositioning to severe deformations. Depending on the allowed warping transformations it may be possible to model regional dilation, contraction, shearing and other changes. Ideally, each non-linear parameter introduced should move additional motion-induced variability out of the registered images and into the mathematical transformations as registration of homologous landmarks improves. To quote Toga and Thompson: 'The

application of warping algorithms is intimately intertwined with the biological question at hand, the source and quality of the data itself and a fundamental understanding of the imposed mathematical transmogrifications' [136].

Warping applications that are particularly demanding include non-linear deformations with large DOF, the use of large or highly sampled volumes as well as identification and extraction of features and characteristics of the data to be utilised in the warping algorithms themselves.

Warping strategies can be *intensity-driven* or *model-driven.* In the former, a measure of intensity similarity is defined between the deforming scan and the target brain. The parameters of the deformation field are successively adjusted until the value of the similarity measure is maximised. Measures based on Mutual Information [50, 141] metrics are suitable for cross-modality matching.

In model-driven strategies, the first step is to build explicit geometric models, representing separate, identifiable anatomical elements on each of the scans to be matched. Anatomical elements are parametrized and matched with their counterparts in the target scan. Higher level structural information ensures biological as well as computational validity. Automatically extracted line models [32] have inherent parametric structure which allows additional geometric features (torsion and curvature, for example) to be included in the matching criterion. *Hybrid models* which incorporate advantages of both model-driven and intensity-driven methods are likely to capitalise on the merits of both. For a detailed examination of non-rigid registration in medical imaging, the reader is referred to [60].

In a Bayesian framework, statistical information about the image is combined with prior information about expected template deformations (the *prior*) to make inferences about the parameters of the deformation field [2]. Such a strategy can potentially be incorporated into intensity-, model- or hybrid-driven methods.

The remainder of this section is allocated to a more detailed examination of a number of these methods which have been applied to the problem of medical image registration. The section concludes with a brief discussion of the potential suitability of such schemes for FMRI motion correction.

4.1.1 Elastic and Viscous Registration Models

Physical continuum models of the deformation maintain the topology and connectivity of the deforming template under many local warps by considering the deforming anatomical image to be embedded in a 3-D physical medium which can be either *elastic* or *viscous*. This assumption provides constraints which make external feature (intensity) forces and internal (smoothing) forces balance. The stipulation that topology be preserved seeks to ensure that structures (such as ventricles) are not broken up or conversely joined (for example, left and right hemispheres) by the warp deformation.

In an elastic medium the displacement resulting from the external intensity-driven forces is assumed to obey the Navier-Stokes equilibrium equations for linear elasticity [135]. Measures of the squared intensity mismatch for F are typically more economical to compute than regional cross-correlation metrics [16, 4, 35] or other matching criteria and are consequently used by most groups working in the field [109, 22, 147]. These metrics are only suitable for mono-modal registration.

In [21], Christensen *et al.* describe a registration approach in which they use a viscous fluid model to control the deformation of the template image to the reference image. The template is modelled as a viscous fluid that flows out to match the study under the control of a Gaussian sensor model. The algorithm is not practical, however, unless a massively parallel computer is available. This prompted Bro-Nielsen and Gramkow to propose a faster algorithm giving a speed

up of at least an order of magnitude [15].

For mono-modal MRI images a Gaussian sensor model (least squares) appears to be an appropriate model of the variation between the template and the study image. The PDE describing the viscous fluid deformation of the template is then solved using scale-space convolution. The authors also go on to argue that the Gaussian filter, used in Demons (see section 4.1.3) instead of the real linear elastic filter is an approximation of the fluid model which could give problems in terms of topology and stability of the fluid model.

4.1.2 Optical Flow

Optical flow [68] is a technique for understanding the content of images by examining the differences caused in images of the same object by motion. Small deformations in the temporal sequences of images are sought and T and S are considered as consecutive time samples in an image sequence $I(x,t)$ where the brightness of brain images are assumed constant in time (mono-modal).

To obtain displacements in directions other than that of the brightness gradient it is necessary to add other constraints. This is due to flow measurement inaccuracy (arising from noise processes and un-modelled changes in the image) and also the *aperture effect* whereby only motion perpendicular to an edge can be computed. These errors can be overcome by smoothing each dimension with a Gaussian filter.

4.1.3 Demons-based Registration Models

A method which has attracted some considerable attention within the medical image registration community over the past few years is the Demons matching

algorithm proposed by Thirion [134]. The main idea, that of having two images to match, is to consider the object's boundaries in one image as a semi-permeable membrane and to let the other image, considered as a deformable grid model, diffuse through these interfaces by the action of effectors situated within the membranes (see figure 4.1).

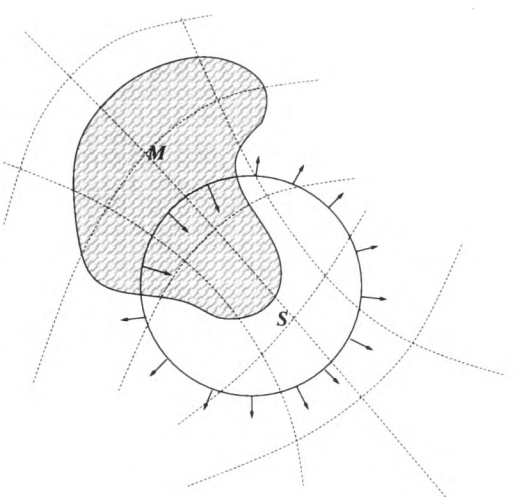


Figure 4.1: Demons algorithm warps a deformable model M onto a static image S by means of effectors situated within the membranes.

Diffusing models rely mainly on the notion of polarity (inside/outside) and attraction relies on distance. For deformable objects with fluid deformations there is not a single definition of an ideal optimal match, but as many definitions as practical applications. Attraction methods include ICP and Snakes [14] algorithms where the emphasis is on distance as opposed to correlation techniques which employ similarity measures but use a local window. It is expensive to search for similar points in the target so optical flow (section 4.1.2) is an alternative.

Instead, the diffusion technique used by the Demons algorithm is based on the concept introduced by Maxwell in the field of thermodynamics. A demon is an effector situated in a point P of the the boundary of an object O. It acts locally to push the model M inside O if the corresponding point in M is labelled

'inside' and outside O if the point is labelled 'outside'.

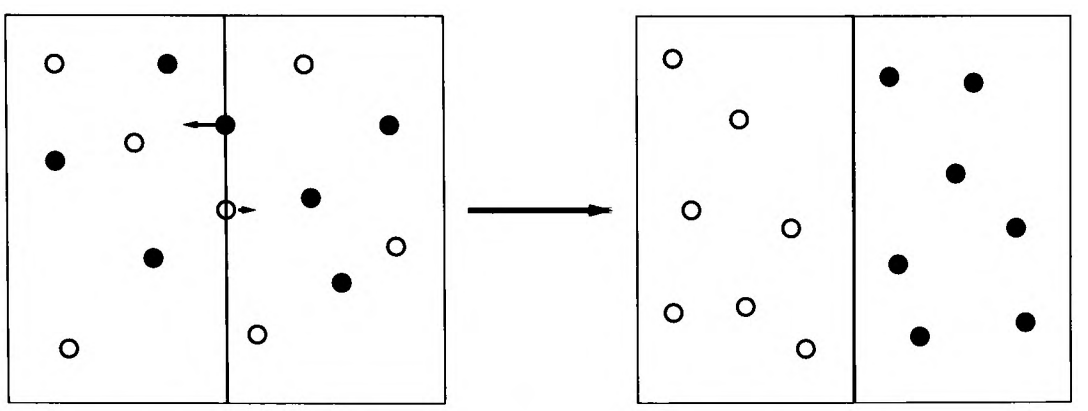
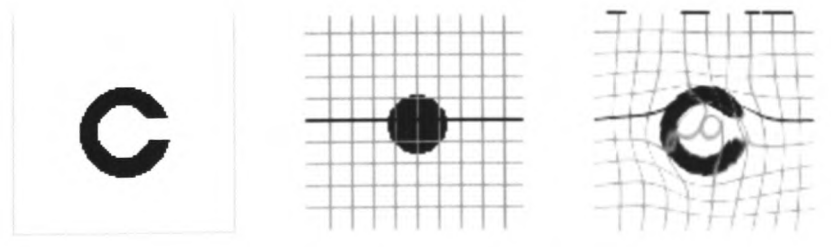


Figure 4.2: The original concept of Demons introduced by Maxwell allowed a membrane to separate two different types of particle onto opposite sides.

A Demons implementation using the full deformable grid (Demons1) was found, by Thirion *et al.* , to be slightly more accurate than one using segmented images (Demons3) but ran three times more slowly including edge extraction. When compared with matching using feature points defined by an anatomist and matching using deformable 3-D crest lines, the three methods were mutually consistent to approximately 3 voxels and all were robust to the choice of reference patient (that is, swapping O and M).

It can be concluded that this level of accuracy is insufficient for the sub-voxel



a. *b.* *c.*

Figure 4.3: Image matching using Demons: *(a)* target image; *(b)* deformable grid model; and *(c)* matching after 30 iterations using $\sigma_{fl} = 3$ and $\sigma_{el} = 3$.

corrections necessary in FMRI and preliminary testing of the Demons algorithm for co-registration of EPI data revealed that it was unable to offer accurate motion corrections, possibly the result also of the intrinsically low resolution of the images.

4.1.4 Multi-modal Warping

While not strictly within the domain of intra-modal motion correction, research into multi-modal registration has yielded a number of interesting model-based techniques. Guimond *et al.* [55] proposed a two-stage iterative scheme which combines a non-rigid warping registration with an intensity transformation step to account for the contrast differences between modalities. Either a mono-functional or bi-functional dependence is assumed to exist between the two modalities (for example, CT and MR). The bi-functional relationship reflects the difference in contrasts between tissues across modalities and allows the intensity s in image S to take one of two values of t in T, determined using a Bayesian scheme.

In both cases the function f is estimated using a Least Trimmed Squares (LTS) estimator and a Re-weighted Least Squares (RLS) scheme [126] to ensure a robust yet information-rich approximation. Once the intensity correction has been performed, the two images are non-rigidly aligned using an algorithm based on the Demons method but also incorporating some aspects of the findings carried out on the SSD minimisation carried out in [19].

4.1.5 Summary of Non-rigid Schemes

The overriding observation which can be made from examining the schemes described above is that because of the level of information required by the al-

gorithms, most are unlikely to be suitable for FMRI motion correction[1]. Mention was made of a preliminary evaluation of the Demon's algorithm (section 4.1.3) on FMRI and the lack of robust or accurate results from this investigation is characteristic of non-rigid intensity-based methods in general.

It has been stated before in section 2.4.2 that landmark-based schemes, while compatible with the non-rigid methods described above, are inappropriate for FMRI where landmarks cannot be detected or identified reliably. Similarly it is not possible, given the resolution of FMRI data, to compute gradients reliably for use in methods such as optical flow or fluid registration schemes.

The most promising avenue of investigation resulting from the survey of non-rigid methods appears to be in the model-based techniques which include constraints which reflect not only the acquisition process leading to the images but also the inherent levels of information contained in them.

A more detailed analysis of the artefacts present in FMRI data is carried out in the next section which, it is hoped, will suggest the general form of a model-based framework for motion correction.

4.2 Limitations of Retrospective Motion Correction

The artefacts arising from residual motion after rigid-body motion correction were illustrated in section 3.3 demonstrating the fundamentally incomplete correction offered by such schemes when applied to real data. It may be unrealistic to expect a retrospective scheme, which does not take these effects into account, to be able to guarantee accurate realignment of images which may vary spa-

[1] At least, given the current state of FMRI.

tially due to activation-related signal changes or because of movement within an inhomogeneous magnetic field.

Furthermore, research has demonstrated that some retrospective motion correction algorithms may in fact create spurious brain activation patterns in the absence of a large level of subject motion [42]. The authors of this work have gone so far as to suggest bypassing the retrospective motion correction stage altogether in the case of moderate yet still appreciable subject motion in order to maintain the integrity of the data. This observation highlights the possibly detrimental effects of a retrospective scheme based purely on image data which in turn motivates the development of a well-integrated scheme in order to deal with artefacts at the source.

More generally it has been noted that so-called 'on-the-fly' motion compensation could reduce acquisition times over a number of medical modalities in addition to reducing the amount of data of poor diagnostic quality and open up new protocols that are currently unavailable due to motion artefacts [64].

More generally, the prospects of FMRI becoming a standard clinical diagnostic tool are greater if the technique can be shown to be efficient and reliable in the clinical setting [95].

4.2.1 Susceptibility

B_0 field inhomogeneities in the scanner cause geometric distortions which are appreciable mainly in the phase encoding direction for EPI [78] and which have been observed to change as the subject moves during acquisition of the EPI data. Retrospective corrections have been proposed which seek to model these deformities as a function of the EPI data and the estimated motion parameters [1].

Un-warping of B_0 field inhomogeneities can be performed before any other

retrospective corrections have been applied to the data. The correction is achieved by registering the EPI to a regularised B_0 field map [72]. Since the distortions are proportional to the strength of the static field, this is likely to be an increasing source of artefact as scanners with higher field strengths are used.

An example of this type of artefact is shown in figure 1.12, where non-linear distortion is present at the front of the brain. For motion correction, however, this would require a field map to be taken at each time point to correct an entire study which is unlikely to be feasible given the time constraints on acquisition already in place. Current research [1, 76, 92, 131] seeks to overcome these problems and in the future motion correction schemes may not be expected to account for gross non-linearities caused by susceptibility.

4.2.2 Slice-to-slice Movement

As EPI volumes are acquired slice-by-slice (see section 2.3.5) there will be a temporal offset between slices from the same volume. Because analysis, specifically fitting a model to each voxel's time-series, assumes that all data within a volume were gathered at an single time point, the data must be adjusted by interpolating along each voxel's time-course so it appears that all images were acquired at the same time.

This correction can be accomplished by Fourier Transforming each voxel's time-course, applying a phase shift to this frequency representation corresponding to the slice in which it was acquired relative to a reference time, and finally reverse-transforming the data [80].

A more common approach is to apply a temporal interpolation over all time-courses in the image data where the amount of interpolation shift is determined by the order in which the slices to which the time-courses belong were acquired.

This is commonly referred to as *slice-timing correction.*

If there has been subject motion during the scanning session, this correction is complicated by the fact that the realigned time-course for a particular voxel may contain data taken from several different slices and at correspondingly asynchronously offset timings. Clearly a comprehensive and accurate correction must take this effect into account.

4.2.3 Interpolation

A point of particular concern is that retrospective motion correction schemes themselves introduce artefacts into the data (section 4.2). A prominent example is that of the interpolation required to reconstruct the realigned data from the motion-corrupted time-series and which may introduce blurring.

This can be seen in the simple example shown in figure 4.4 where a combination of coarse sampling and standard interpolation methods leads to a poor reconstruction of a simple signal profile. The profile shown (a simple step edge) is common in FMRI where the expected profile would be a discrete transition between intensity values at the boundary between two different tissue types.

To first order, the effects should manifest themselves in the temporal domain as a linear combination of periodic functions of linear combinations of the realignment parameters [53]. The spatial analogue to this effect will be linear combinations of sinc kernels resulting in ringing patterns at high frequencies emanating from areas with large spatial derivatives.

An example of interpolation artefact isolated by ICA is shown in figure 4.5.

Use of more accurate (and commensurately slower) interpolation schemes such as the sinc-kernel interpolation employed as the final estimation and realignment stages of some motion correction algorithms, including SPM99, can reduce such

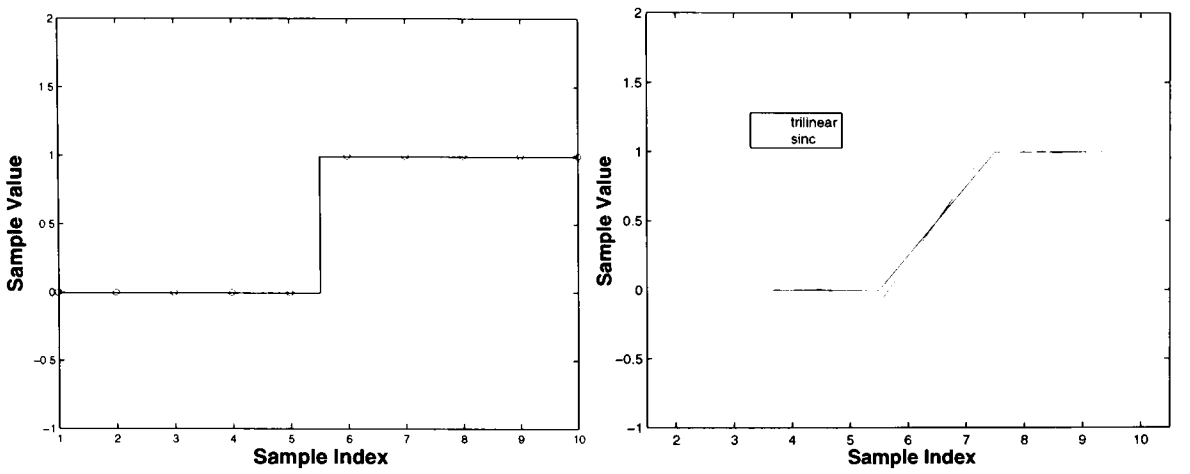


Figure 4.4: Example of interpolation artefact: interpolated values, shown on the right, were generated in MATLAB by re-sampling a discrete step edge with trilinear and sinc kernels respectively using only the data points marked with a circle in the left-hand plot.

errors [108] but not completely eliminate them. Work has been done on a retrospective correction for interpolation errors introduced by motion correction [53]. A periodic function of the estimated displacements is used to remove interpolation errors on a voxel-by-voxel basis but this involves making further assumptions about the data and may in fact lead to a further loss in information. In general, having to apply an additional adjustment to correct for artefacts introduced by an earlier realignment should suggest that the realignment stage needs to be redesigned so as not to distort the images further.

The goal of the work described in this chapter and the next is the development of a scheme which, by maintaining a constant reference frame between subject and scanner gradients, would reduce the need for retrospective realignment and the associated errors identified above.

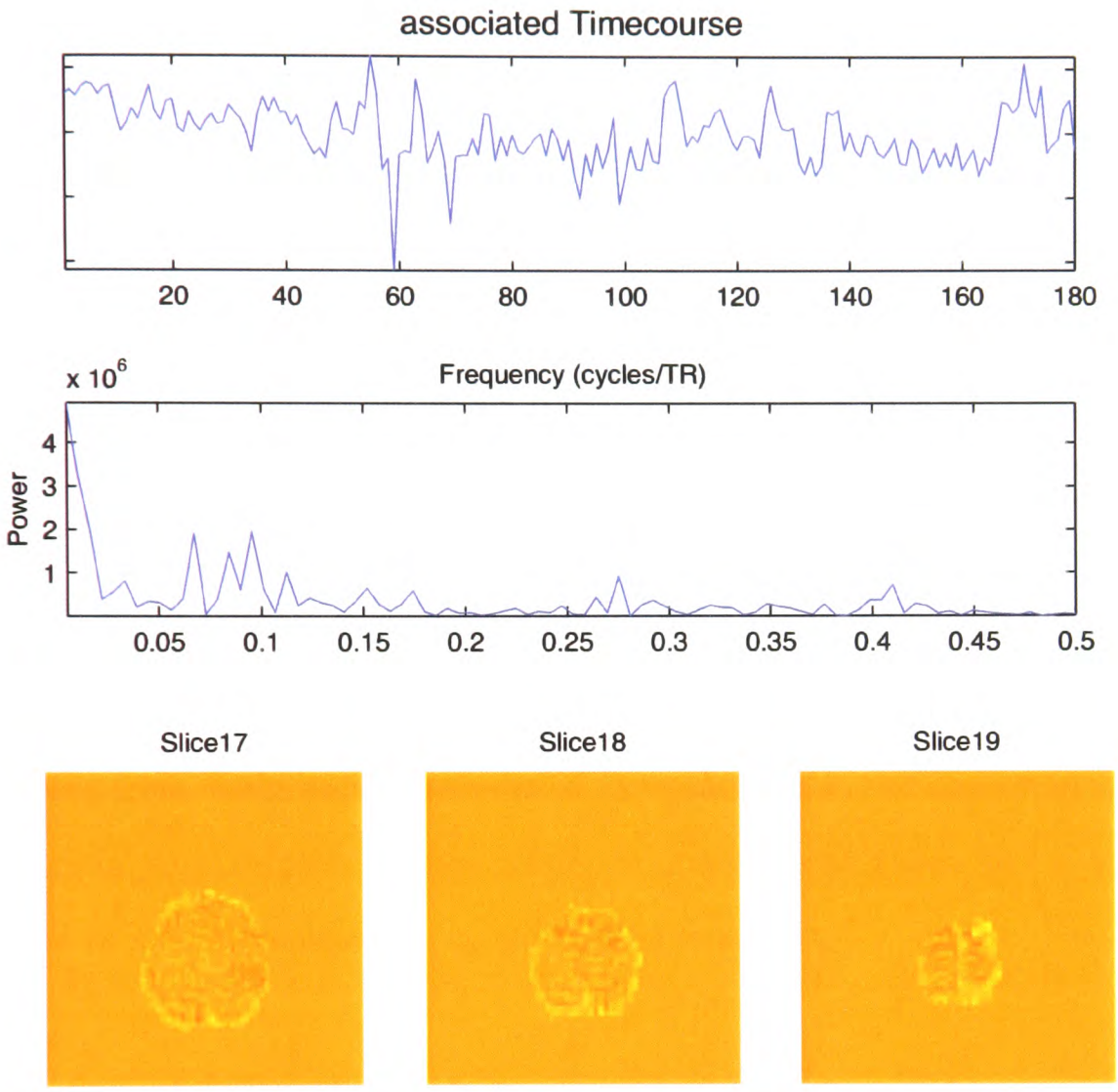


Figure 4.5: IC time-course and selected spatial maps showing interpolation artefact with spatial manifestation around the edge of the brain.

4.3 Prospective Methods

Following on from the work using orbital navigator echoes, prospective correction schemes to date have taken advantage of the basic property of ONEs, namely that they provide a measure of subject motion *during* the scan which can be used to adjust the gradients/ RF pulse applied to subsequent acquisitions to compensate for motion. This is in contrast to intensity-based schemes which can offer only a measure of motion artefact once the entire data-set has been collected if typical

repeat times are to maintained

In the case of ONEs, it is then possible for this information to be acted on immediately and a correction performed before the next scan takes place. Clearly, if such a scheme is to co-exist with established acquisition methods, the estimation of motion from the navigators and subsequent processing and feedback of the information to the system (including any necessary hardware setting updates) must all take place in the time between scans, usually much less than 3 seconds.

4.3.1 Through-plane Correction

Early research sought to correct for through-plane motion by using navigators sensitised to superior/ inferior (S/I) motion of the subject [95].

Starting from the notation introduced in equation 2.2, the change in phase of the navigator with respect to some reference phasor (with phase $\Psi(\theta - \alpha)$) after a rotation of α and translation (x_0, y_0) can be written:

$$\Delta\Psi(\theta) = \Psi(\theta) - \Psi(\theta - \alpha) = k_p(x_0 \cos\theta + y_0 \sin\theta) \tag{4.1}$$

Recall that the angular rotation can simply be obtained from the magnitude of the displaced navigator. The translation can be found using the orthogonality constraint on sin and cos terms. For example:

$$x_0 = \frac{1}{\pi k_p} \int_0^{2\pi} \Delta\Psi(\theta) \cos\theta \mathrm{d}\theta \tag{4.2}$$

In [95] a navigator was acquired and processed in under 100 ms for a single-slice study to give an estimation of Δz with respect to a baseline initial meas-

urement which was then used to adjust the pulse-sequence controller before the next EPI acquisition was triggered. Δz was taken as the value of distance along the z axis, d which minimised:

$$l(d) = \sum_{z=W/2}^{W/2-1} |S_d(z) - S_R(z)|^2 \tag{4.3}$$

with respect to the reference navigator echo $S_R(z)$ over a window of consideration W. Studies using a motor-controlled phantom which moved between 5 mm and 25 mm in fixed increments between navigator acquisitions showed that the data had a standard deviation almost an order of magnitude smaller after correction using this scheme.

4.3.2 Rotational Motion

While compensating for some of the through-plane errors in acquisition, further developments of this scheme are required to correct for in-plane rotations and translations so as to account for all forms of rigid-body subject motion. If this can be achieved it may be possible to eliminate the need for retrospective rigid-body motion estimation and realignment interpolations.

Working on the basis that a single ONE is able to give a measurement of view-to-view rotation and 2-D translational motion of an object, a scheme has been proposed whereby three orthogonal ONEs were applied in advance of each EPI acquisition [94, 142]. In-plane translation is estimated prospectively, along with the other parameters, but is only applied retrospectively, yielding a saving of 50 ms on the total time for each correction at what the authors see as a negligible degradation of the quality of the data.

It might be argued that prospective translational corrections should reduce not only interpolation error but also other artefacts associated with subject motion as described earlier in this chapter. The scheme described in [142] takes 160 ms to estimate and update the rotational and through-plane translational errors. The final implementation requires 520 ms for the prospective correction between scans to allow for two estimation steps (the rotational parameters are corrected after the first 160 ms cycle in order to improve accuracy of motion estimation in the second cycle) and data transfer to the scanner before the next EPI acquisition.

While giving accurate results (both sub-millimetre translations and sub-degree rotations were obtained), the time involved will require that standard experiments specifying a TR of around 3 seconds or less will have to be modified in order to take advantage of this scheme and still allow the usual excitation times needed by the RF field. In addition the level of correction offered by this scheme may still be too coarse to provide the levels of accuracy required.

4.3.3 Octant Navigators

Recent work has demonstrated a new navigator echo sequence which is able to gauge rigid-body rotation and translation in a single short read-out [139]. The navigator works by tracing out the outline of an octant on the surface of a sphere in k-space. A full 6 DOF (3 rotation and 3 translation) read-out can be acquired in a few milliseconds allowing even quite rapid subject motions to be characterised. The additional overhead required by this scheme is to perform a navigator pre-map over a range of rotations and translations to serve as a reference for the later navigators. This takes a few seconds at the start of the scan and gives measurements which are optimal for the particular subject.

4.3.4 Implementation Details

For the speeds quoted by prospective methods to be achieved and the data integration to take place in the brief window between acquisitions, consideration must also be given to the hardware implementation. In particular the schemes described in sections 4.3.1 and 4.3.2 use a real-time array processor originally developed for fluoroscopic imaging. In general most schemes of this nature require specialised hardware in order to achieve the data processing and transfer times necessary.

4.4 Comparison of Approaches

In the case of non-rigid registration, it is possible to accommodate a much greater level of flexibility into the motion correction algorithm. Experience has shown, however, that this increased freedom comes at the cost of reliable and repeatable results. As outlined in section 4.2, FMRI data-sets do not offer the information-rich and detailed images on which such schemes rely in order to provide accurate corrections.

A secondary drawback to a fully non-rigid scheme is that typically the algorithms take considerably longer to run than their affine or rigid-body counterparts introduced in Chapters 2 and 3. This limits their usefulness in a clinical setting where a large volume of data must be processed in a limited time.

In contrast, the prospective methods outlined show great promise in the future for FMRI motion correction. Set against this potential is the relative infancy of such approaches, not to mention the need for additional hardware, which must be overcome they can be adopted by the FMRI community at large. It is also widely acknowledged that, based on the demonstrated accuracy of existing

navigator-based corrections, any prospective scheme will still require a retrospective intensity-based correction stage to achieve the required sub-voxel levels of accuracy needed for FMRI analysis. This hypothesis is supported by the levels of accuracy provided by such schemes and reported in section 2.3.3.

From the examination of non-rigid schemes contained in this chapter, it can be concluded that the best results for image registration are typically achieved where a robust model for the acquisition process is used. Such a scheme should also be biased to the strengths of the data, rather than to areas where the data are unable to support detailed or complex calculations. In the specific case of FMRI it seems intuitive that a novel robust framework should seek to obtain its information from the volume, rather than the resolution, of the data.

In Chapter 5 a detailed model for motion is proposed which takes into account the likely magnitude and characteristics of motion in FMRI. Based on this model, a correction scheme is developed and tested and comparison is made with existing rigid-body correction schemes such as MCFLIRT.

Chapter 5

Spatio-temporal Registration

Motion correction schemes which incorporate exclusively spatial information and apply a registration algorithm to realign FMRI data fail to address the temporal characteristics of motion artefacts observed in the data. Furthermore, any rigid-body realignment will inevitably incorporate incorrect assumptions regarding the nature of subject motion in FMRI, specifically that motion takes place discretely between whole-volume acquisitions.

The most prominent omission from such rigid-body schemes is the failure to take into account the discrete slice sampling within each volume which can allow a range of independent movements throughout the acquisition of a single volume. Failure to model the process of independent slice movement can lead to inaccurate reconstruction of the data, which will further result in incorrect statistical inferences regarding the intensity and spatial manifestation of any neurological response being observed.

This chapter begins by investigating the artefacts which can arise if rigid-body assumptions are made about subject motion in FMRI. It is demonstrated that in the case of many general types of subject motion, the use of rigid-body motion correction can actually result in a degradation in the data, rather than

an improvement.

Based on these observations and by applying some of the techniques reviewed in the previous chapter, a novel spatio-temporal model of subject motion is introduced. This representation seeks to model accurately the spatial transforms describing subject motion in FMRI along with associated slice-timing information, and to use this to provide a more precise correction than any other retrospective method to date.

This model leads to an integrated spatio-temporal correction scheme which is implemented and tested in Chapter 6. The results of this testing are presented and show that it is possible to remove significantly more motion artefacts than previous schemes have been able to do.

5.1 Intra-volume Temporal Artefacts

When the basic concepts of FMRI were introduced in Chapter 1, it was explained that Echo Planar Images are typically acquired in a slice-sequential manner, that is, all the voxels falling within a single slice of scanner space are captured before the voxels in the next slice. While acquiring more than one slice at a time is possible in some MRI applications, it is currently impractical for FMRI because the resulting signal-to-noise ratio is too low to be usable.

Single-shot EPI (as opposed to multi-shot, interleaved or segmented EPI) is the EPI sequence that is typically used in FMRI. The 'single shot' refers to the fact that each slice is reconstructed from a single RF excitation. Although it is possible to acquire several slices with single shot — one shot for each slice — the FMRI data-sets acquired to date are nearly always single-shot multi-slice EPI.

Echo Volumetric Imaging is the full extension of EPI to 3-D which means that

the full volumetric image is acquired in a single shot. It is extremely difficult to carry out in practice [23] and results in a very low resolution image (currently, about $32 \times 32 \times 16$ voxels) with large levels of distortion due to an even lower SNR than is typical of multi-slice EPI.

Given that SNR considerations limit FMRI to stacked-slice acquisitions in practice, the problem of temporal offsets within a volume due to the successive acquisition of slices remains a significant confound to motion correction. Previous attempts to correct FMRI data for artefacts introduced during acquisition have considered spatial realignment and slice-timing correction as two distinct and separate stages in the processing chain as shown in figure 5.1.

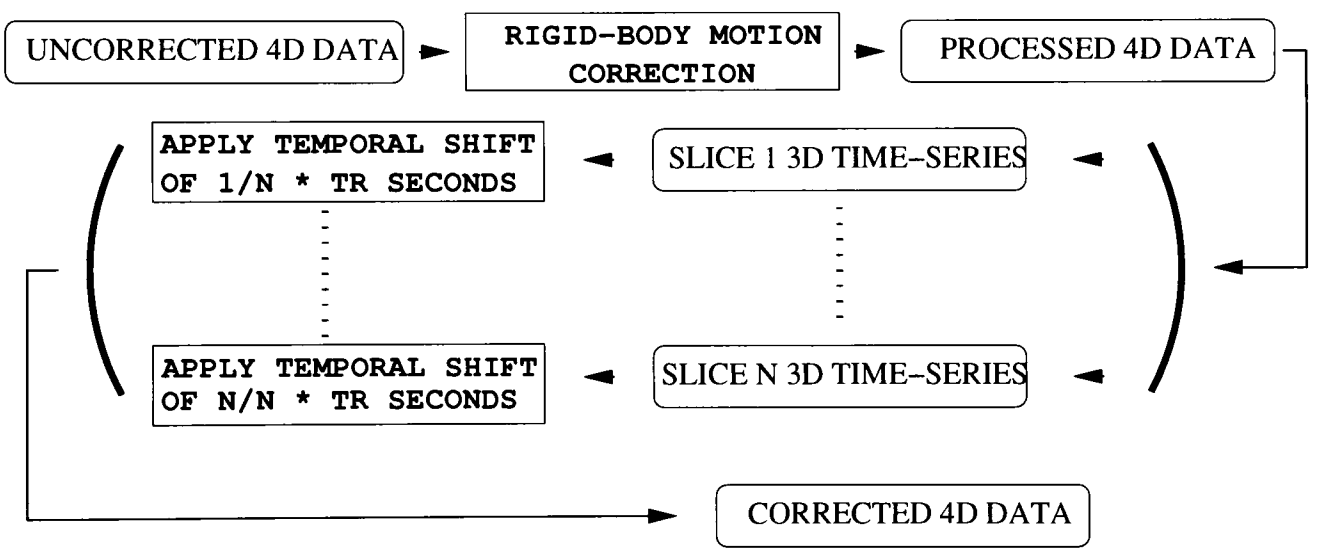


Figure 5.1: Schematic of the separate application of spatial and temporal realignment of FMRI data. Note that in this arrangement, motion correction is applied before slice-timing. The reverse ordering is discussed in section 5.1.1.

Performing slice-timing correction means, ideally, that the image reflects a true 'snapshot' of the object at a discrete point in time, rather than as a volume of slices acquired sequentially. The slice-timing-corrected data should more accurately reflect the spatio-temporal relationship between voxels throughout the volume. Because the EPI data-sets under consideration have typically been ac-

quired slice-by-slice rather than by using a volumetric approach, the temporal distribution of the slices within each volume must be taken into account if correct and meaningful conclusions are to be drawn about the spatio-temporal relationships within the data.

5.1.1 Existing Approaches

While the need for slice-timing is widely acknowledged, these corrections are not always performed in practice on FMRI data. Applying the two corrections separately is convenient, not least because it facilitates the use of existing tools such as MCFLIRT and SPM for rigid-body motion correction along with separate temporal interpolation of each voxel time-course to re-shift slice-timings. There are fundamental errors in the assumptions underlying this distinction, however, regardless of the order in which the two steps are performed.

Clearly, if no subject motion has occurred, it is sufficient simply to apply slice-timing correction as a series of temporal interpolations over each voxel time-course in turn, where the amount of shift is proportional to the temporal offset associated with the slice containing the voxel being considered, shown in figure 5.1. A complete lack of subject motion is unlikely to occur in real data, however, so the interaction between motion and acquisition delays must be modelled in order to correct fully for the resulting artefacts in the data.

Assuming that motion correction is carried out before any temporal corrections, data which may not correspond to acquisition at a consistent point in time will be co-registered. If slice-timing correction is applied after the initial realignment, the corrected images will contain data from several discrete sample times within individual slices. This is because, in the general case of through-plane motion, spatial registration will realign the data so that intensity values from

individual slices in scanner space are distributed across several slice locations in the corrected data. This is illustrated in figure 5.2.

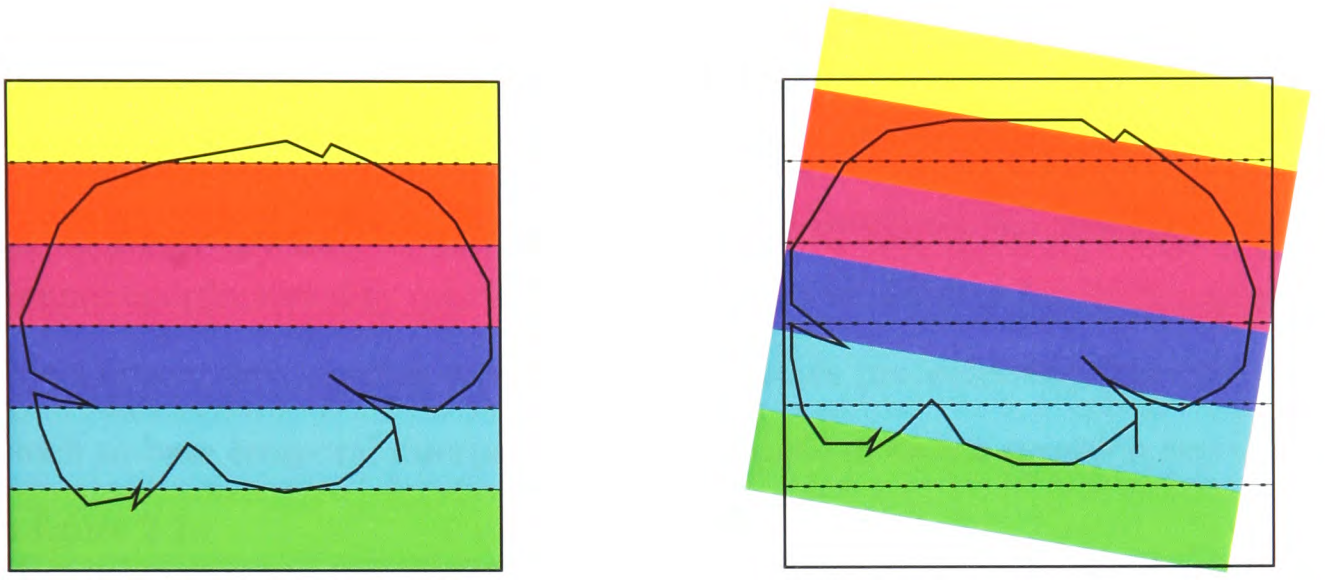


Figure 5.2: In this figure the coloured bars, fixed to the same reference frame as the brain, indicate the acquisition timings associated within each slice. Even in the case of rigid-body motion, spatial realignment (in this case correcting a simple pitch of the head) will lead to a breakdown in the correlation between slice location used to determine slice-timing correction (marked with dotted lines) and associated acquisition timings (coloured bars).

Specifically, if rigid-body realignment is performed, subsequent slice-timing will make the incorrect assumption that data within individual slices will have been acquired at the same time-point. In this situation, it is necessary to keep a record of the slice in which the data were originally acquired and then apply the appropriate timing correction, a step which is usually omitted.

It might therefore seem obvious that temporal re-sampling should be carried out before motion correction. An obstacle to such a re-sampling is that in order to carry out slice-timing correction by temporal interpolation of a particular voxel, the time-course of that voxel must be known. If the subject has moved, there is no guarantee that a voxel in object space will be in constant alignment with a voxel in scanner coordinates. This creates a cyclic problem where motion correction is needed in order to determine slice-timing before motion correction. Recognising

this inter-dependency is by no means original, but prior to this work there have been no attempts to utilise knowledge about the acquisition sequence in order to correct the situation.

The situation is worse when considering a voxel on an intensity boundary (for example on the interface between two tissue types or on the perimeter of an activating region), since in such a case the voxel described by a particular set of scanner coordinates may rapidly switch between two different intensity regions in object space, thus creating a physically implausible (uncorrected) time-course on which to base temporal interpolation. An example of such a situation was shown in figure 2.1.

In general, if a rigid-body model is assumed, the motion correction stage will also ignore the fact that there may not be a parallel correspondence[1] between the slices from different volumes if motion has occurred during or between activations, such as the example in figure 5.4. Thus a rigid-body spatial realignment will inevitably attempt to compute pair-wise voxel comparisons on data which originates from different spatial locations in the subject's brain. It is also possible that the slices will remain unaligned, even when the volumetric optimisation has reached a minimum. In conclusion, movement throughout a scan will lead to different displacements in individual slices. This is ignored by volumetric corrections, such as MCLFIRT, which assume a rigid-body transformation over the entire volume.

These observations show that the separate application of slice-timing and motion correction cannot accurately account for motion artefacts in FMRI. For this reason, the work described in this chapter proposes an integrated approach to these two corrections that is able to cope with the potential spatial non-linearities

[1]That is, it is no longer correct to assume that the relative movement between slices is consistent across different volumes.

in the data.

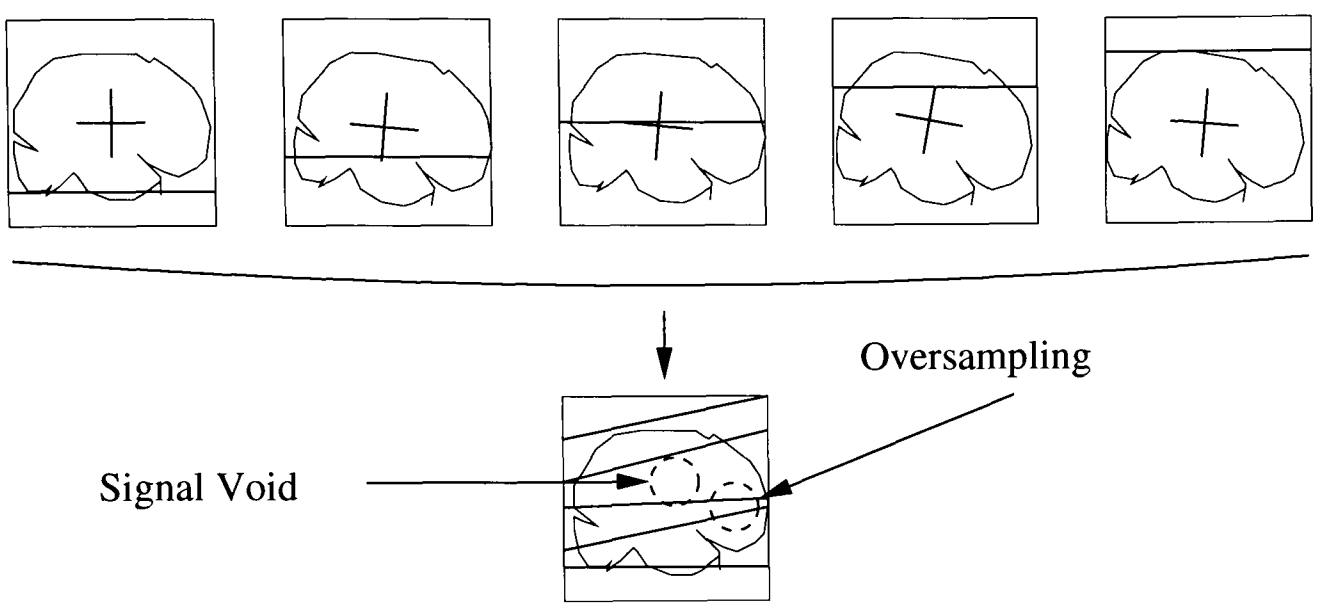


Figure 5.3: Movement during the acquisition of a volume (in this case a simple nodding motion is depicted) may cause acquisition redundancy and voids.

5.1.2 Non-linearities

If the subject has moved during acquisition of that volume, a quite likely possibility, as previously noted, is that there will be redundancy at some locations yet sparseness and/or voids at other points of the volume. This is due to a particular location in the object having moved sufficiently quickly between slice acquisitions to have been captured either more than once or not at all, as shown in figure 5.3. This observation is not new in the FMRI literature: incorrect assumptions about spatial distribution of slices have already been identified as the cause of errors in the realignment of FMRI data [111, 27].

At a more basic model level, because subject motion may occur more markedly during some volume acquisitions than in others, the initial motion correction stage may be trying to infer an affine relationship between two volumes which do not have the same overall morphology (such as the example shown in figure 5.4).

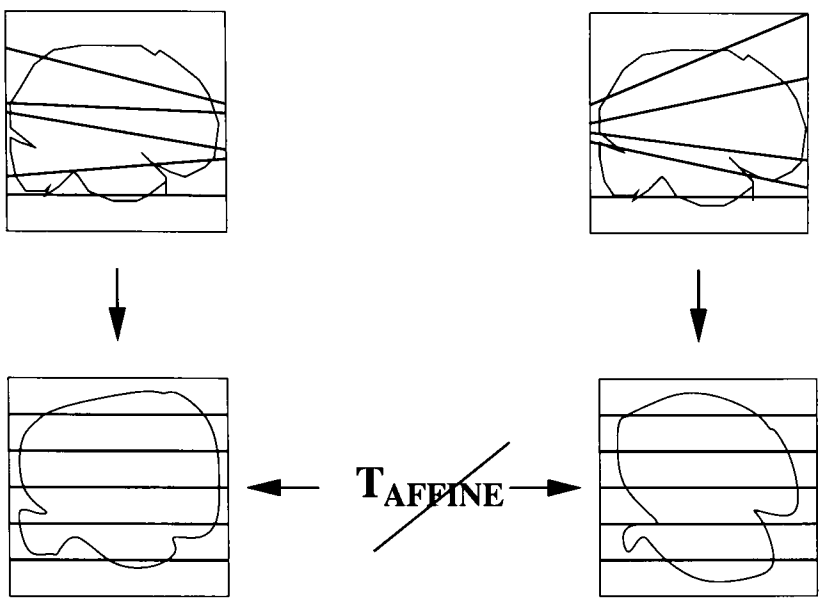


Figure 5.4: Local distribution of slices within different volumes may demand a non-rigid solution to the motion correction problem.

Another flaw in applying standard intensity-based volumetric registration in this situation is that each volume will contain a unique spatio-temporal distribution of voxels. While all the voxels within each scanner slice are temporally aligned, it is not possible to say before motion estimation which physical locations in object space these voxels relate to or whether this object-to-scanner correspondence is consistent across all volumes.

A method that registers individual slices to a single anatomical volume has been proposed in [86] but this does not take into account the timing issues and requires a cross-modal registration. The latter condition is particularly confounding as any 2-D to 3-D registration of this kind is bound to be poorly constrained (the most pathological example of this would be to trying to match a circle to a unique plane within a sphere) and even more so in a cross-modal application where one of the modalities is low resolution EPI, with all its inherent distortions.

Estimation of slice-timing and motion inevitably relies on voxel intensities acquired through a relatively low resolution modality (including activation-related

signal changes of interest), so a spatio-temporal scheme offers the additional advantage that it may be possible to combine information over a large number of images to give more robust conclusions than those normally available from pair-wise realignment of EPI volumes.

5.1.3 In-plane Motion

To confirm that separate rigid-body motion correction and slice-timing are insufficient for the somewhat easier case of no slice overlap, a data-set was generated from a high-resolution EPI image containing $192 \times 256 \times 128$ voxels, representing a FOV of $192 \times 256 \times 163.8$ millimetres, to which independent planar rotations around the central z-axis, shown in figure 5.5 were applied using trilinear interpolation to each set of 6 slices. This grouping was chosen to reflect subsequent down-sampling into the resolution expected in FMRI data.

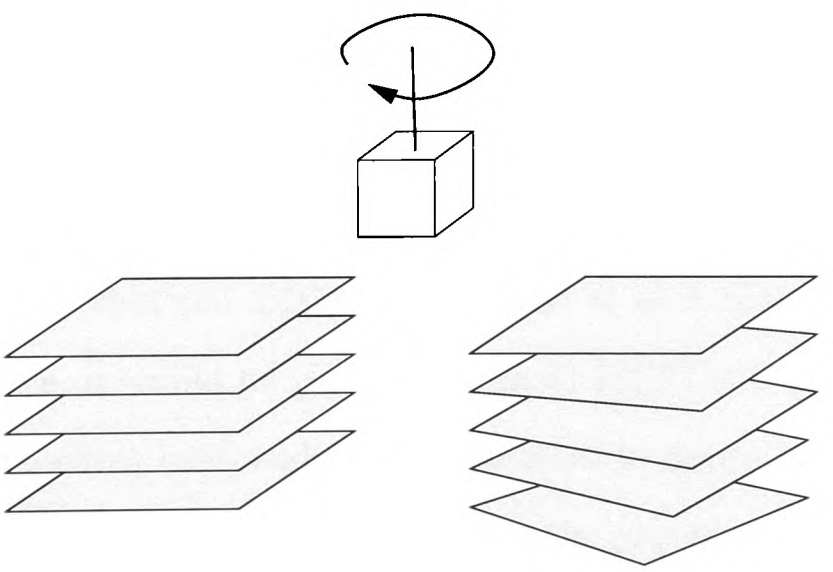


Figure 5.5: Diagrammatic representation of the shaking motion (rotation about the z-axis only) introduced into the first artificial data-set.

The range of rotation was $+/-2$ degrees over any single volume and the maximum rotational deviation from the original orientation was 4 degrees. The

variation of the z-rotation is shown in figure 5.6. Once the motion had been added, the images were down-sampled (using neighbourhood averaging) to a more common resolution of $64 \times 64 \times 21$ voxels in order to minimise the magnitude of any interpolation artefacts arising from the post-translation re-sampling. A representative volume, after down-sampling, is shown in figure 5.7.

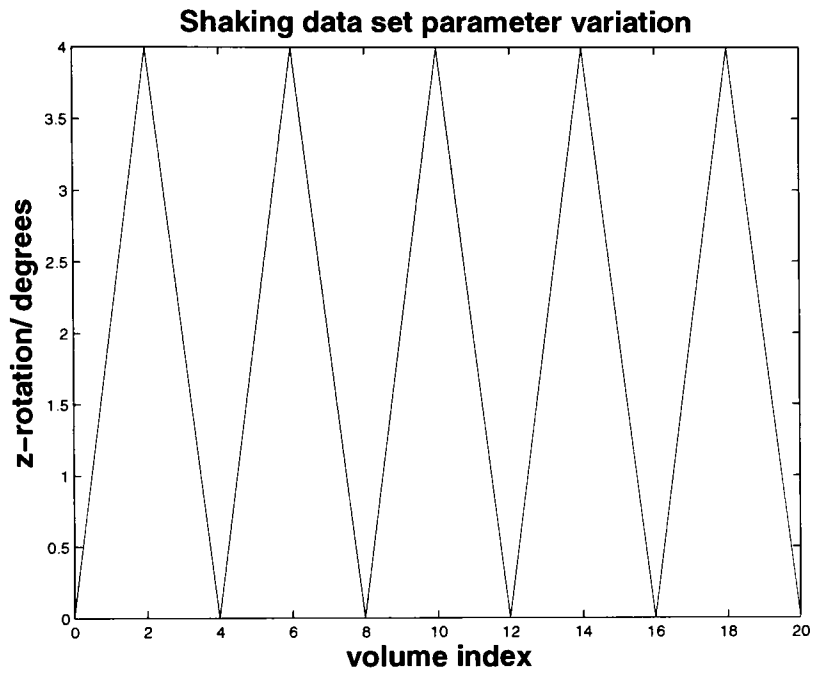


Figure 5.6: Plot representing the variation of z-rotation across the 20 volume time-series containing shaking motion. Each time-point corresponds to a unique volume index.

If it were known that the motion was limited at each time to a single slice, as in the case here, it would be possible to apply MCFLIRT motion correction by processing the slices independently as depicted in figure 5.1. The problem with this method is that it can only work when the precise nature of the subject motion is known, either by accurate observation of the subject during imaging or through some pre-registration step. Both methods would be incompatible with a repeatable, non-interventional algorithm and without such informed data manipulation a standard volumetric motion correction will result in an inadequate correction.

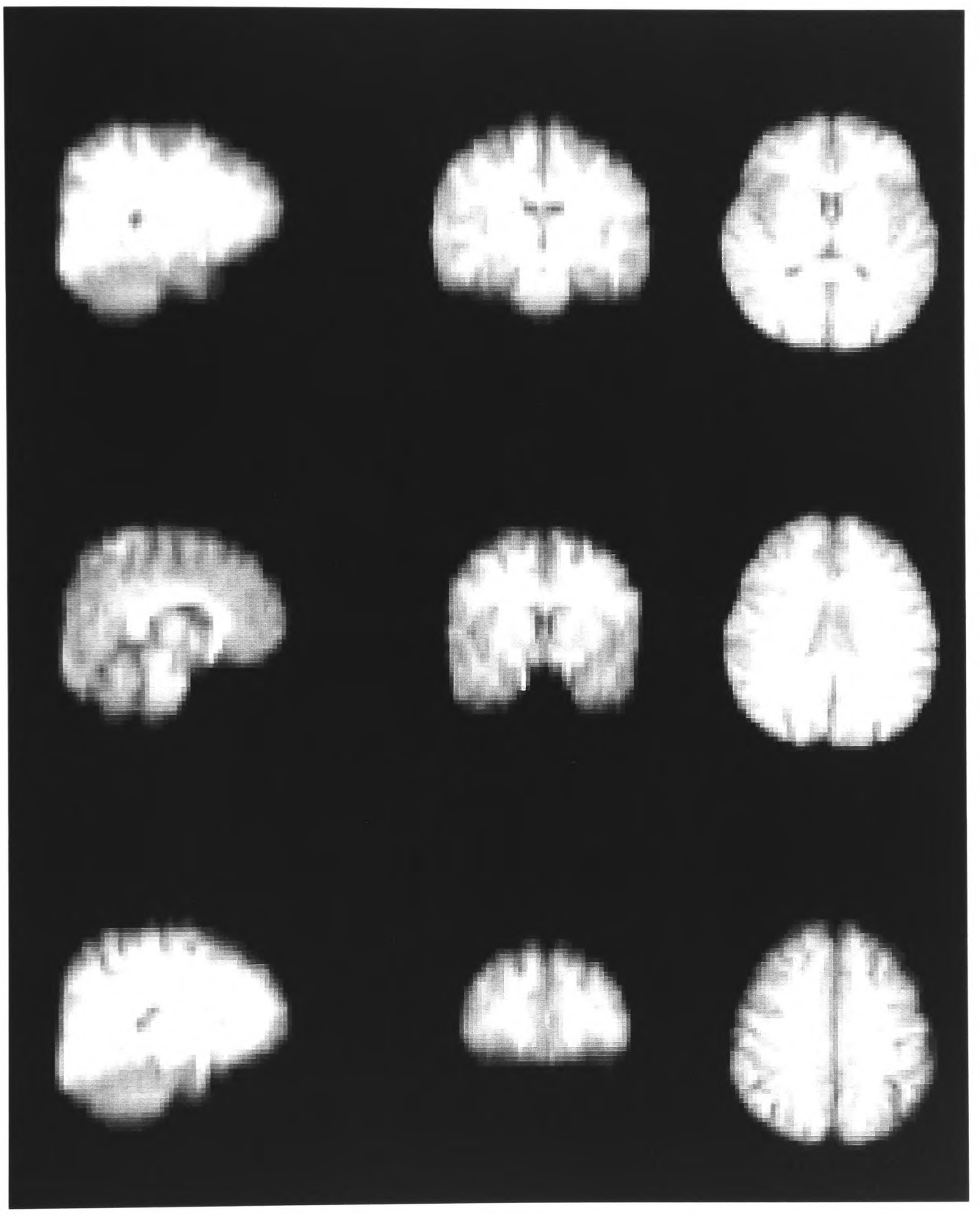


Figure 5.7: Sample volume from an artificial time-series containing in-plane movement only. This figure is included to confirm the quality of the artificial data compared with real FMRI data and to illustrate that the effects of independent-slice movement are too subtle to be identified by manual inspection.

Slice-timing correction was implemented as a 1-D sinc interpolation applied to each time-course separately, with an offset proportional to the relative position of the slice containing the time-course. Using MCFLIRT either before, after or without this temporal correction, the residual error in the data after any correction, measured as the median average residual variance (MARE) for each voxel's time-course, can be compared as shown in figure 5.8.

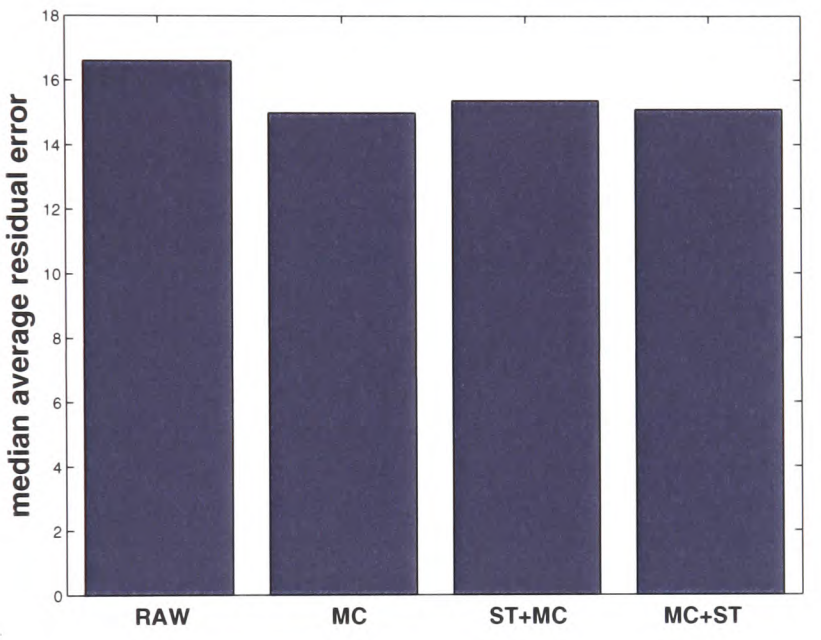


Figure 5.8: Comparison of relative correction errors, measured as median average residual variance, of separate spatial and temporal schemes when applied to data-sets which are known to contain only intra-slice movement within individual volumes. *From left to right:* Uncorrected data (RAW), MCFLIRT motion correction (MC), slice-timing correction then MCFLIRT (ST+MC), MCFLIRT then slice-timing correction (MC+ST)

It is clear that the incorrect rigid-body model for the type of motion contained in the data can lead to a poor correction. Further, due to the incorrect modelling of the spatio-temporal relationship, there is no advantage in applying slice-timing correction, which in this case actually further degrades the data. This is an important result, as work on FMRI motion correction to date has assumed that the volumetric model is approximate, but not inaccurate.

5.1.4 Through-plane Motion

In order to demonstrate the potentially large inaccuracies resulting from the use of separate spatial and temporal realignment of FMRI data, a sample data-set, similar to that described for in-plane motion, was constructed according to a known motion design. The images simulated subject motion in a 'nodding' pattern where each volume contained rotational motion about the y-axis (left to right) projected through the centre[2] of the image and ranging from -2 to $+2$ degrees over the whole volume as depicted in figure 5.9. The variation of x-rotation with time-point is plotted in figure 5.10.

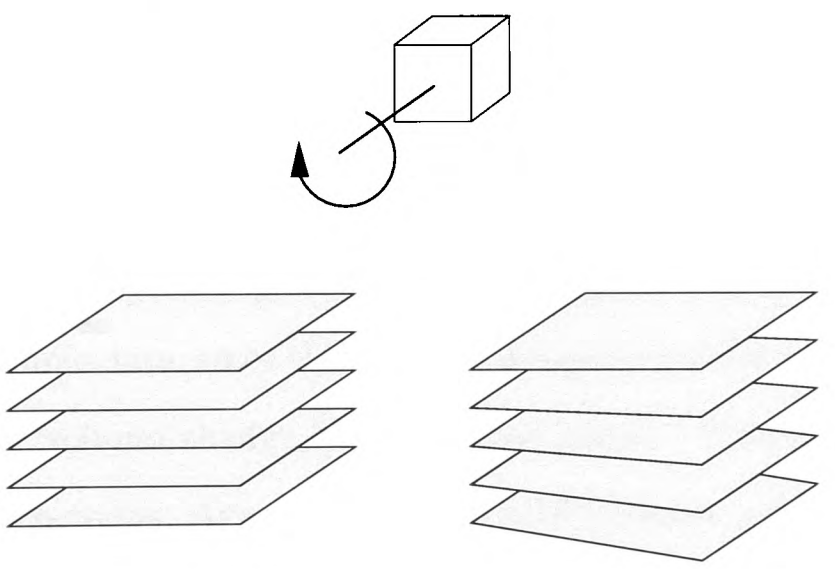


Figure 5.9: Diagrammatic representation of the nodding motion introduced into the second artificial data-set.

This motion was applied incrementally to each set of 6 slices to simulate the intra-volume motion found in many real data-sets. In other words, if the motion over a volume was determined to be 2 degrees, the first 6 slices were pitched forward by 0.095 degrees, the second set by 0.19 degrees and so on up to a rotation of 2 degrees for the final (21^{st}) set of 6 slices.

[2]Taken to be the centre of gravity computed over the intensity values of the template image.

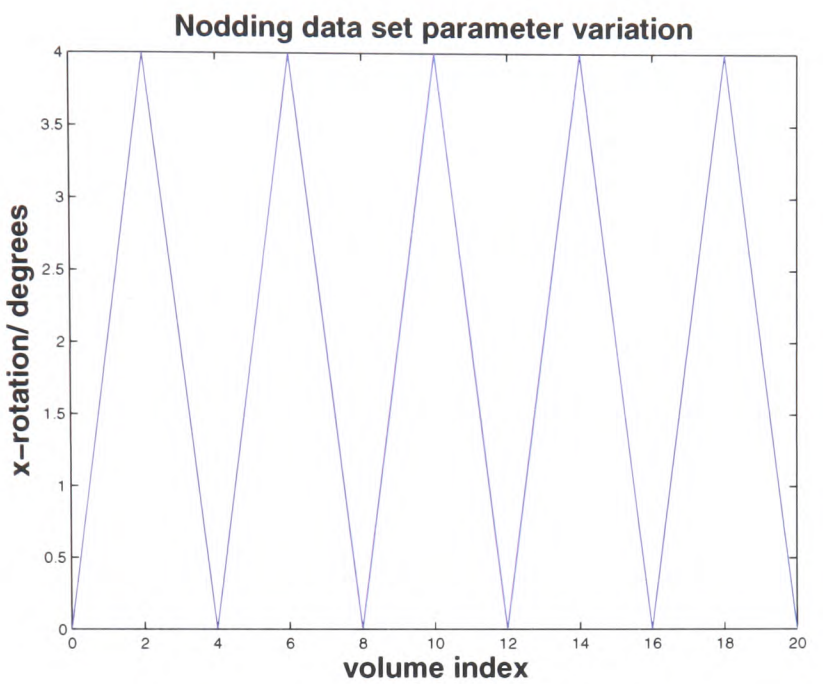


Figure 5.10: Plot representing the variation of x-rotation across the 20 volume time-series containing nodding motion. As in figure 5.6, each time-point maps to a unique volume index.

Even if the nature of the motion were known accurately, it would no longer be possible to adopt the parallel spatial correction approach that could be used for the previous example data, since the number and source of the voxels intersecting each slice of the volume change between time-points. Figure 5.11 shows the effects of such corrections. Again using the MARE measure, results show that for through-plane motion not only does rigid-body motion correction and slice-timing fail to correct the motion-related distortions but the results of the corrections are actually *worse* than no correction at all. This is due in part to the averaging effect that a rigid-body estimation will have when the slice motion consists of perturbations around some baseline pose, as is the case in this data-set. Motions which occur at a low level in the first few slices of the volume will tend to be underestimated while the largest motions which have been applied to the later slices will be under-estimates as the rigid-body model cannot account for the range of transformations contained in the volume as a whole.

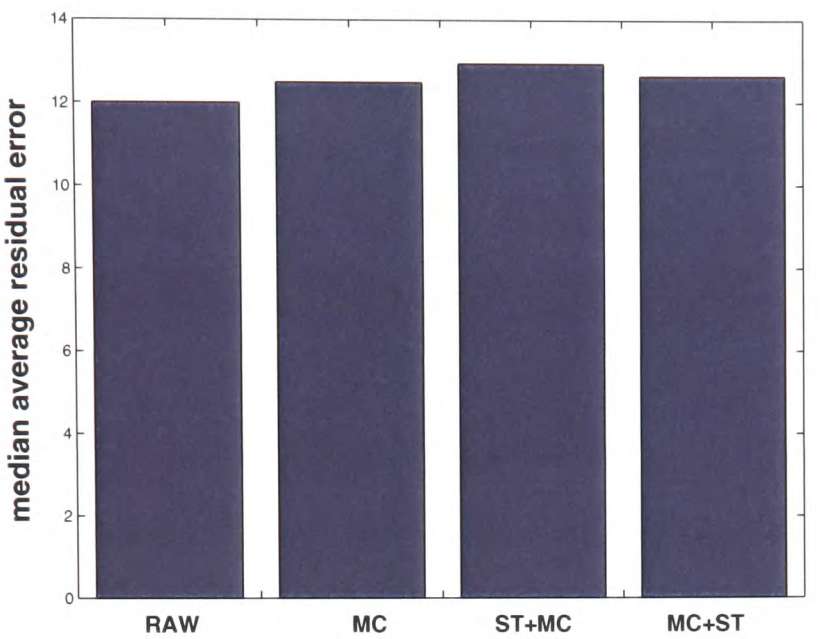


Figure 5.11: Comparison of relative accuracies, measured as median average residual variance, of separate spatial and temporal corrections when applied to data known to contain inter-slice movement within individual volumes. *From left to right*: Uncorrected data (RAW), MCFLIRT motion correction (MC), slice-timing correction then MCFLIRT (ST+MC), MCFLIRT then slice-timing correction (MC+ST).

Clearly, both in the case of through-plane and in the earlier example of in-plane motion, standard motion correction techniques using rigid-body registration and simple voxel-wise temporal interpolation for slice-timing cannot accurately compensate for the types of motion typically encountered in FMRI. While MCFLIRT cannot account for motion at the discrete slice level, it may prove possible to refine the rigid-body estimates and as such they may form a sensible initialisation for a non-rigid scheme. As mentioned before, it is critical that a non-rigid scheme does not attempt to infer information, such as highly-localised non-linear motion, which cannot be directly and reliably extracted from the images. Based on the earlier observations regarding the need for integrated spatio-temporal motion correction, and the clear demonstration of the shortcomings of existing techniques, a new model for subject motion in FMRI is required.

In the next section, such a model is developed and used to drive the devel-

opment of an integrated approach to motion correction which is intended to be sufficiently flexible and robust to accurately remove a wide range of motion artefacts. Before moving on to this stage, however, it is instructive to examine the behaviour of separate motion correction and slice-timing on data-sets exhibiting much more subtle levels of motion, which are typical of studies with cooperative subjects or controls.

These new data-sets were produced using either the nodding or shaking motion described above, but with +/− the maximum rotation applied to each volume, that is, every second volume is identical (figures 5.12 and 5.13). Because the temporal characteristic of the motion alternates between equal-sized positive and negative values, slice-timing will now tend to average these values over time and may give the impression of a good correction. Even so, it is impossible to predict which combination of slice-timing and motion correction will yield more accurate results, further reinforcing the arguments in section 5.1.1 regarding the incorrect application of these two separate steps.

The results, shown in figures 5.14 and 5.15, depict this erratic performance. The plots also reveal that without the averaging effect of temporal interpolation, motion correction alone will still lead to a degradation of the images.

5.2 Spatio-temporal Model

The new model described in this section is intended to bypass the shortcomings of the primarily sequential methods described above. The acquisition process is modelled as a set of discrete time-points within each volume, one corresponding to each slice, so as to allow changes in the orientation of the head for each slice.

A practical approximation to this process is to express the orientation of each

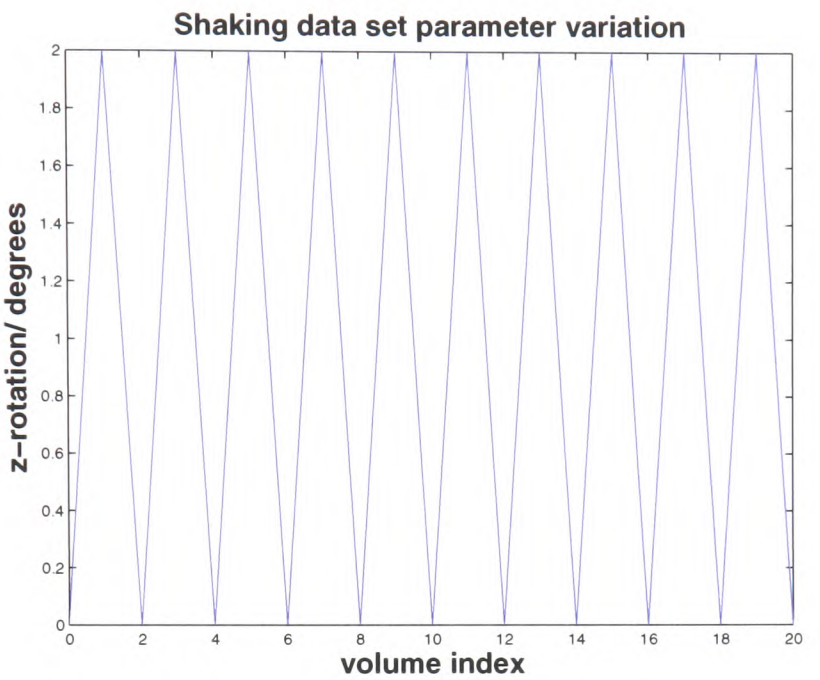


Figure 5.12: Plot representing the variation of z-rotation across the 20 volume time-series containing shaking motion at a lower level than before.

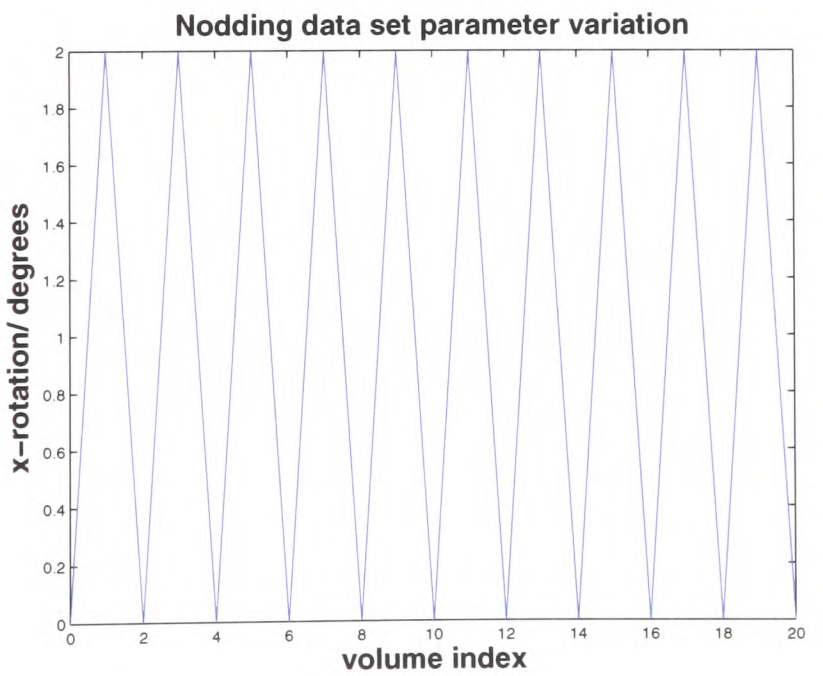


Figure 5.13: Plot representing the variation of x-rotation across the 20 volume time-series containing nodding motion at a lower level than before.

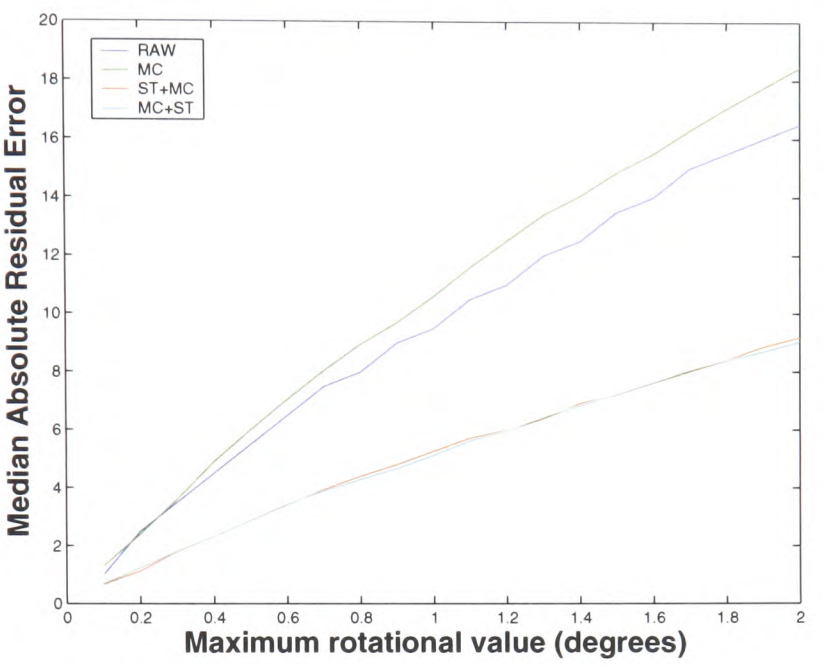


Figure 5.14: Plots of the median average residual variance for a shaking motion design show the effect of correction with separate slice-timing and rigid-body motion correction on images containing very low levels of motion. Notation corresponds to uncorrected data (RAW), MCFLIRT motion correction (MC), slice-timing correction then MCFLIRT (ST+MC), MCFLIRT then slice-timing correction (MC+ST).

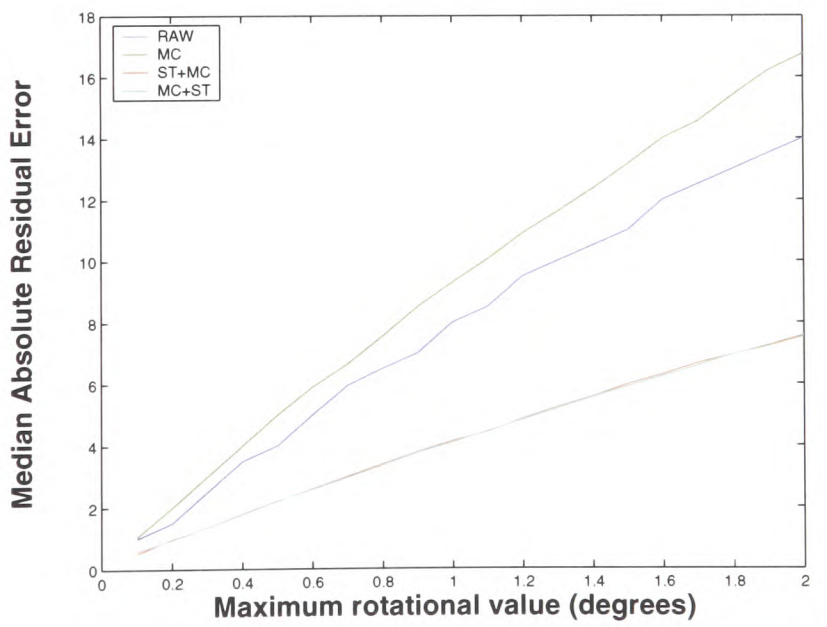


Figure 5.15: Plots of the median average residual variance for a nodding motion design show the effect of correction with separate slice-timing and rigid-body motion correction on images containing very low levels of motion. Notation corresponds to Uncorrected data (RAW), MCFLIRT motion correction (MC), slice-timing correction then MCFLIRT (ST+MC), MCFLIRT then slice-timing correction (MC+ST).

slice as a fraction of the estimated gross motion contained in the volume. This is done by estimating the motion between two volumes and decomposing this motion into a set of evenly-spaced rotations about a common axis [116]. The assumption made at this stage is that the rotation is in one sense (that is, there is no sign change in the derivative). It should be noted that the associated increase in DOF is only a total of 6 additional parameters for the whole time-series. Thus the assumption of an even distribution of slices along some postulated rotational axis is the key to keeping the subsequent optimisation problem manageable. This is discussed in more detail below. Due to the increase in the number of DOF for the registration problem, however, it is necessary to devise a method which is able to cope efficiently with the permitted flexibility in the data and which applies constraints and some global cost invariant over the data as a whole.

The proposed method can generally be divided into 3 separate sections: the initialisation stage, which transforms the images from volumetric data into distributed, discrete slices; the cost function optimisation, with its requisite spatial interpolation choices, which serves to refine the initial image pose estimates by exploiting spatio-temporal similarities in the data; and, finally, the temporal interpolation which implements spatially-integrated slice-timing corrections and includes the estimation of intensity values associated with voxels which are deemed not to have been acquired as a result of inter-slice movement.

These stages are discussed in detail in the following sections, which conclude with a summary of the spatio-temporal model in the form of pseudo-code (section 6.1). This is followed by testing of the implemented model in section 6.2.

5.2.1 Slice Pose Estimation

Given two volumes which have been acquired one after the other in an FMRI time-series, it is possible to estimate the total motion of the subject's head within the first volume V_n, as the rigid-body transformation $\boldsymbol{T_n}$ which relates V_n to the second volume, V_{n+1}. This is the approach adopted by MCFLIRT and other rigid-body motion correction schemes. Although this estimate can only be an approximation in the case of an independent-slice model, because the movement is intra-scan rather than volumetric, it provides a useful initial approximation which can then be refined progressively throughout a subsequent optimisation process.

It is assumed that T_n is smoothly varying (sudden intra-scan movements within the volume are possible; but in this first study the model will avoid this additional level of complexity in the interests of testing the underlying hypothesis of significant inter-slice motion) so that the transformation can be decomposed into a rotation θ about some common axis, $\mathbf{q} = q_x\mathbf{i} + q_y\mathbf{j} + q_z\mathbf{k}$:

$$\boldsymbol{T_n} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} q_x^2 \text{vers}\,\theta + \cos\theta & q_y q_x \text{vers}\,\theta - q_z \sin\theta & q_z q_x \text{vers}\,\theta + q_y \sin\theta & t_x \\ q_x q_y \text{vers}\,\theta + q_z \sin\theta & q_y^2 \text{vers}\,\theta + \cos\theta & q_z q_y \text{vers}\,\theta - q_x \sin\theta & t_y \\ q_x q_z \text{vers}\,\theta - q_y \sin\theta & q_y q_z \text{vers}\,\theta + q_x \sin\theta & q_z^2 \text{vers}\,\theta + \cos\theta & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.1)$$

where $\text{vers}\,\theta = (1 - \cos\theta)$, the versine of θ.

This formulation, or 'screw decomposition', makes it possible to construct a set of incremental rotations about $\mathbf{q}$ by allowing the angle of rotation to range as $\theta_s = s.\Delta\theta$ where $\Delta\theta = \frac{\theta}{S}$, where s is the current slice number and S is the number of slices. Similarly the translational component $\mathbf{t} = t_x\mathbf{i} + t_y\mathbf{j} + t_z\mathbf{k}$ can be applied in increasing fractions up to the full (rigid-body) value.

These slice-based transformations can be written as $\boldsymbol{T_{n,s}}$ where $s = 1, .., S$. By applying each incremental transformation to the corresponding plane in scanner space it is possible to form an independent-slice representation of the volume. This is illustrated in figure 5.16. This decomposition provides a more realistic interpretation of the data in each slice with respect to the subject's head than the standard stacked-slice model [86, 110].

Furthermore, because the distribution of slices is determined by a limited number of parameters (specifically $\mathbf{q}$, θ and $\mathbf{t}$ - all of which are derived from the

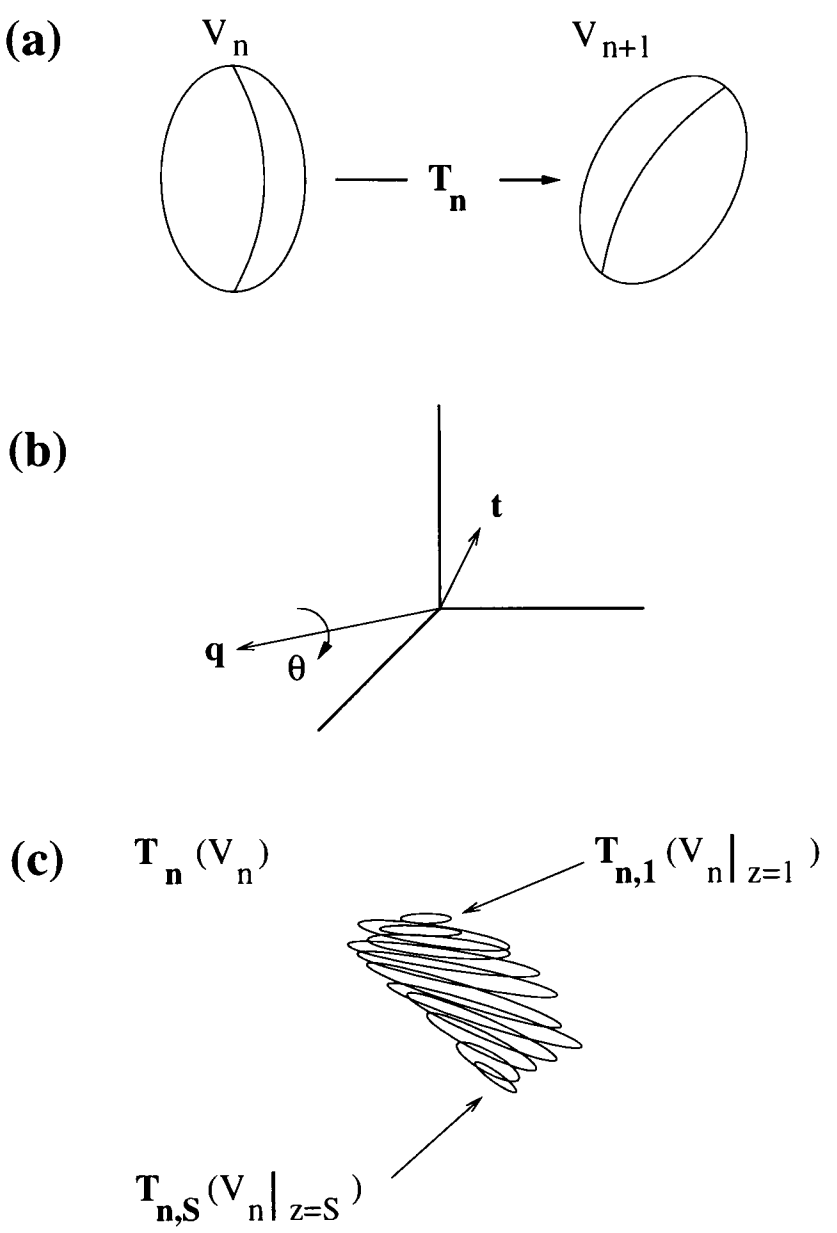


Figure 5.16: By expressing the rigid-body transformation $\boldsymbol{T_n}$ relating volume V_n to volume V_{n+1}, shown in **(a)**, as a rotation θ about a single axis **q** and a translation **t**, depicted in **(b)**, it is possible to determine a set of transformations $\boldsymbol{T_{n,s}}$ using equation 5.1 which describe the local movement of n individual slices. These can be applied to each slice of volume V_n, written $V_n|_{z=s}$ where $s = 1, .., S$, to give volume $\boldsymbol{T_n}(V_n)$ shown in **(c)**, where each slice is transformed by progressively greater fractions of the original transformation, $\boldsymbol{T_n}$. This provides a more accurate model of the data acquired on a slice-by-slice basis within each volume.

rigid-body matrix $\boldsymbol{T_n}$) it is possible to update the orientation of the slices by varying only $\boldsymbol{T_n}$, thus limiting the complexity of a correction scheme based on this model. This is an important aspect of the model, given the dimensionality of the spatio-temporal data and the number of sample points.

An important corollary to this observation is that by altering $\boldsymbol{T_n}$, and thus the slice distribution within the test volume V_n, the distribution of slices within the adjacent reference volume V_{n+1} will also be affected. This may be thought of as an 'accordion' effect where the first slice of the test volume and the final slice of the reference volume are constrained while the midpoint slice is varied, thus altering the spacing of slices within each volume. An illustration of this relationship between volumes is given in figure 5.17.

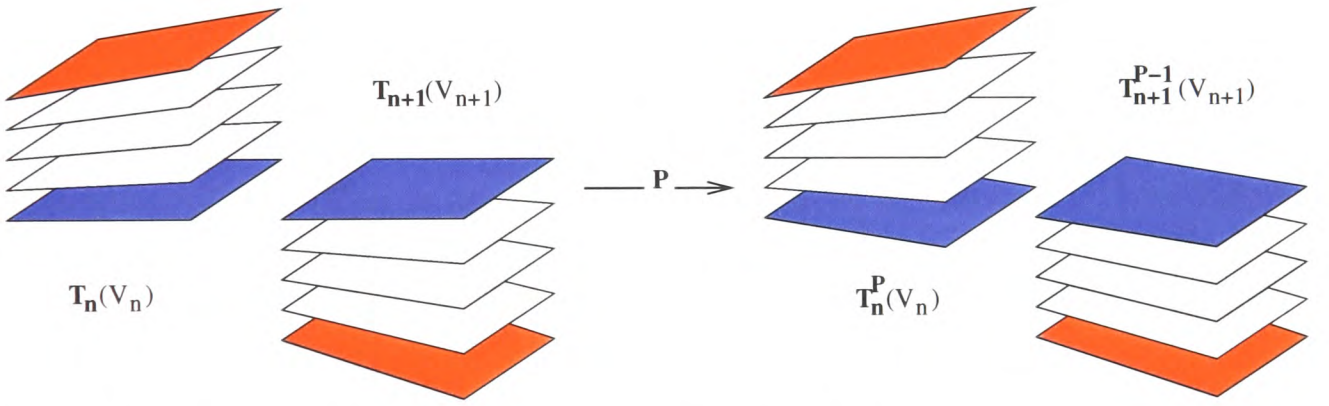


Figure 5.17: At each optimisation step for a volume V_n, a perturbation $\boldsymbol{P}$ is applied to $\boldsymbol{T_n}$ and the new slice distribution $\boldsymbol{T_{n,s}^P}(V_n)$ tested. Throughout this optimisation step, the end slices of the volume pair $[V_n, V_{n+1}]$ shown in red are kept fixed. In effect what $\boldsymbol{P}$ does is to vary the mid-point slice between these two volumes, shown in blue. Thus, for every perturbation applied to V_n it is necessary to apply the inverse $\boldsymbol{P^{-1}}$ to $\boldsymbol{T_{n+1}}$ (giving $\boldsymbol{T_{n+1}^{P^{-1}}}(V_{n+1})$) to maintain the constraint on the end (red) slices.

For example, if $\boldsymbol{T_n}$ relates slice 1 of V_n to slice 1 of V_{n+1} and $\boldsymbol{T_{n+1}}$ relates slice 1 of V_{n+1} to slice 1 of V_{n+2}, then if slice 1 of V_{n+1} moves by $\boldsymbol{P}$, the transformation $\boldsymbol{T_n}$ followed by $\boldsymbol{P}$ will map slice 1 of V_n to the (now perturbed) slice 1 of V_{n+1}. If the updated rigid-body transformation associated with V_n is written

$$\boldsymbol{T_n^P} = \boldsymbol{P} * \boldsymbol{T_n} \tag{5.2}$$

where $\boldsymbol{P}$ describes the parameter updates determined by optimisation (the *perturbation matrix*), then the rigid-body transformation T_{n+1} describing the orientation of the next volume V_{n+1} (which is based on its relative position with respect to V_n) must be adjusted. To understand this, consider the relationship between the newly perturbed slice 1 of V_{n+1} and its original orientation in the example above, which equals $\boldsymbol{P^{-1}}$. Therefore, the updated $\boldsymbol{T_{n+1}}$ becomes $\boldsymbol{P^{-1}}$ followed by the original $\boldsymbol{T_{n+1}}$. According to the relationship between V_n and V_{n+1} described in section 5.2.1, this update can be written:

$$\boldsymbol{T_{n+1}^{P^{-1}}} = \boldsymbol{T_{n+1}} * \boldsymbol{P^{-1}}. \tag{5.3}$$

This ensures that the product $\boldsymbol{T_{n+1}} * \boldsymbol{T_n}$ is the same, independent of the application of $\boldsymbol{P}$, which is necessary as the product of the two transformations describes the mapping between slice 1 of V_n and slice 1 of V_{n+2}, neither of which have moved.

By performing this slice-wise decomposition for every image in the series, it is possible to obtain an initial approximation to the data as a more realistic representation of the subject's head in scanner coordinates, rather than a volume corrupted by a single volumetric transformation (illustrated in figures 5.4 and 5.3). This initial pose is obtained by decomposing the volumetric estimates given by MCFLIRT using the method described above. While this decomposition is still approximate, it is now possible to refine the orientation of the slices within each volume using the optimisation procedure described in the next section.

5.2.2 Cost Function

Having constructed the initialised distributed-slice time-series by applying the model from section 5.2.1 to each volume, the next stage in correcting the data is to construct a cost function which attains a minimum when the 4-D time-series is aligned.

The assumption underlying this stage, as with conventional intensity-based rigid-body schemes designed to work on FMRI data, is that there will be a small level of intensity variability restricted to a small number of voxels. These voxels are assumed to correspond either to regions of the subject's brain which have activated under the stimulus of the experiment or those areas which have been affected by an acquisition artefact. The intensity of the remaining voxels should remain broadly constant in object coordinates so the correction should seek a realignment over the 4-D data which minimises this variance.

A simple candidate for the cost function is Least Squares which, for each voxel's estimated time-course over a time-series of length N, can be computed and summed over all voxel locations in the 3-D volume to give:

$$\sum_{\mathbf{x}} \sum_{n=1}^{N} (\boldsymbol{T_n}(V_n(\mathbf{x})) - \overline{V(\mathbf{x})})^2 + \text{const.} \tag{5.4}$$

While conventional Least Squares seeks to measure the squared difference in corresponding voxel intensity values between a target volume V_n and a single reference volume V_{n+1} (see Appendix III), this measure characterises the squared difference between the slice-transformed $\boldsymbol{T_n}(V_n)$ acquired at time n and the average of the volumes in the time-series.

This metric is intended to reduce the likelihood of local movement patterns dominating the evaluation of a minimum. An example of such a situation would

be where there is no appreciable motion throughout two adjacent volumes but where those two volumes are significantly offset from the rest of the time-series. By generating a reference image from the entire remainder of the time-series, there is a much smaller possibility of the presence of an adjacent (similarly misaligned) volume driving the evaluation of the cost function.

Note that this calculation differs from the use of a mean template image in MCFLIRT (section 3.2.1) in that the intensity values which are input to the cost function are those which have been transformed to their 'true' positions by means of the slice decomposition described in section 5.2.1.

At this stage, it is not necessary to adjust the intensity values to reflect the temporal offset associated with each slice. This is because the steady state areas which are driving the cost function towards an optimal value are assumed to be invariant, regardless of the order of acquisition.

Once the initial slice distribution has been determined via the process described in section 5.2.1, a target volume V_n(initially the first in the time-series) is selected for realignment. In order to quantify the cost, the variance of the remaining volumes is computed for comparison. Note that because the orientation of slices within V_{n+1} also depends on the pose of V_n, it is necessary to omit V_{n+1} from the grouped reference measure.

For a time-series of length N this measure, computed over $N-2$ volumes, can be written:

$$Var_{N-2} = \frac{SumSq_{N-2} - (N-2)\bar{E}^2_{N-2}}{N-3} \tag{5.5}$$

where the sum of squared intensity values can be computed as

$$SumSq_{N-2} = \sum_{\forall i \in N : i \neq n, n+1} (\boldsymbol{T_{i,s}}(V_i))^2 \tag{5.6}$$

where $\boldsymbol{T_{i,s}}(V_i)$ denotes the application of $\boldsymbol{T_{i,s}}$, $s = 1, .., S$, to the volume V_i. The mean is then simply

$$\bar{E}_{N-2} = \frac{1}{N-2} \sum_{\forall i \in N : i \neq n, n+1} \boldsymbol{T_{i,s}}(V_i). \tag{5.7}$$

Note that Var_{N-2} will remain invariant for any given test volume V_n, as $T_{n,s}$ are changed. Thus in order to compute the cost update under a new slice distribution of the test volume, $\boldsymbol{T_{n,s}}(V_n)$, the new mean (incorporating the test volume) will be

$$\bar{E}_{N-1} = \frac{(N-2) * E_{N-2} + \boldsymbol{T_{n,s}}(V_n)}{N-1}, \tag{5.8}$$

and the sum of squared intensity values can be similarly modified to include the test volume:

$$SumSq_{N-1} = SumSq_{N-2} + (\boldsymbol{T_{n,s}}(V_n))^2, \tag{5.9}$$

and so by substituting equations 5.8 and 5.9 into equation 5.5, the new cost including the proposed slice distribution of the test volume, becomes

$$Var_{N-1} = \frac{SumSq_{N-1} - (N-1)\bar{E}^2_{N-1}}{N-2}. \tag{5.10}$$

Finally the spatial mean, which is used by the optimisation process described in section 5.2.3, is computed:

$$\overline{Var_{N-1}} = \frac{1}{x_{max}y_{max}z_{max}} \sum_{\forall x,y,z \in \mathbf{x}} Var_{N-1}(\mathbf{x}) \tag{5.11}$$

5.2.3 Optimisation Method

In section 5.2.1 a method was presented for calculating the slice orientation of a volume V_n directly from $\boldsymbol{T_n}$, the rigid-body transformation relating it to an adjacent reference volume V_{n+1}. In order for the optimal orientation to be determined, it is necessary to vary $\boldsymbol{T_n}$ and thus the slice distribution within V_n (and V_{n+1}) until the cost function given by equation 5.11 reaches a minimum.

In order to arrive efficiently at a robust and accurate solution, an optimisation process similar to that used by MCFLIRT is implemented. Golden section search is looped over each of the 6 rigid-body parameters in turn (see section 3.1.3) until a minimum is found.

An optional second stage is offered where simple gradient descent is carried out, again on each parameter in turn with a smaller tolerance than that allowed for the golden section search. An evaluation of the optimisation is presented in section 6.2.2 and demonstrates the accuracy obtained both with and without the additional gradient descent search.

At each iteration of the optimisation process, the updated transformation parameter is incorporated into $\boldsymbol{T_n}$, giving a new transformation $\boldsymbol{T_n^P}$ (equation 5.2). This transformation is decomposed in turn, according to equation 5.1 and applied to the test volume to give an updated volume using the new slice distribution, written $\boldsymbol{T_{n,s}^P}(V_n)$.

This new volume is substituted for the previous volume decomposition, $\boldsymbol{T_{n,s}}(V_n)$, into equations 5.8 and 5.9 to allow the updated cost to be computed from equation 5.11. It is this cost which is minimised by the optimisation process.

Searching in each of the 6 parameter directions halts when a specified tolerance has been reached. Once the minimum cost has been found, it is necessary to update the orientation of the next volume V_{n+1}, from which the individual slice matrices describing its internal orientation are computed (equation 5.3). The overall scheme then substitutes V_{n+1} for V_n as the test volume and proceeds to optimise $\boldsymbol{T_{n+1}}(V_{n+1})$.

For a time-series of length N, the realignment has approximately $6N$ degrees of freedom: one for each of the rigid-body parameters which control the movement of the end slice in each volume. The slice distribution throughout the volume is a fixed distribution which is calculated from the transformation between the first and last slices in the volume, as shown in figure 5.16.

5.2.4 Interpolation Method

As each realignment in the optimisation process will impact on the relative movement between volumes, slice orientation is updated at each stage of the optimisation process. The data may be non-uniform after this slice redistribution (see figure 5.4), so a 3-D regularisation [74] is used to determine values over the time-course for a given voxel location.

An important observation is that because of the potential sparseness in the data (it is now assumed that there is a much larger degree of temporal independence than in standard motion correction algorithms because individual volumes are no longer considered to represent samples taken at a single, consistent time-point) interpolated values are always derived from the original data. This condition is enforced so as to avoid compounding any early approximations with later interpolation.

The current intensity value in the volume V_n at location $\mathbf{x}$, $I_n(\mathbf{x})$, is generated

'on the fly' by applying the inverse slice transformation

$$\mathbf{x}' = \boldsymbol{T}_{n,s}^{-1} * \mathbf{x} \tag{5.12}$$

and performing a sinc interpolation on the location $\mathbf{x}'$ in the original data given by this back-calculation to determine the value of $I_n(\mathbf{x}')$.

As it is possible, under extreme subject motion, for a slice in scanner space to contain data from almost any location in the head, it is necessary to try all the possible slice transformations associated with V_n, $V_{n,s}, s = 1, .., S$, and take the value of $\mathbf{x}'$ where the slice number of $V_{n,s}$ matches the z-coordinate of $\mathbf{x}$. This step is illustrated in figure 5.18. By way of illustration, consider two slice matrices: $V_{n,X}$ maps a particular voxel from slice X to slice Z, while $V_{n,Y}$ maps another voxel from slice Y to slice Z. Assuming that the two transformations are not the same, it becomes apparent that while applying $V_{n,X}^{-1}$ to a particular voxel in slice Z will map the voxel back to slice X, the same transformation will not map the other voxel back to slice Y.

In the event that the subject has moved such that there are multiple acquisitions of a particular voxel, each related to a separate slice matrix, then a simple weighted sum of contributing voxel values is applied [124]. It is believed that much of the accuracy in this scheme lies in the quality of the interpolation as the motion itself is assumed to be small once the initial slice distribution of the data (described in section 5.2.1) has been carried out.

5.2.5 Temporal Re-sampling

In addition to the intermediate interpolation described in the previous section, care must also be taken when performing the temporal re-sampling applied to

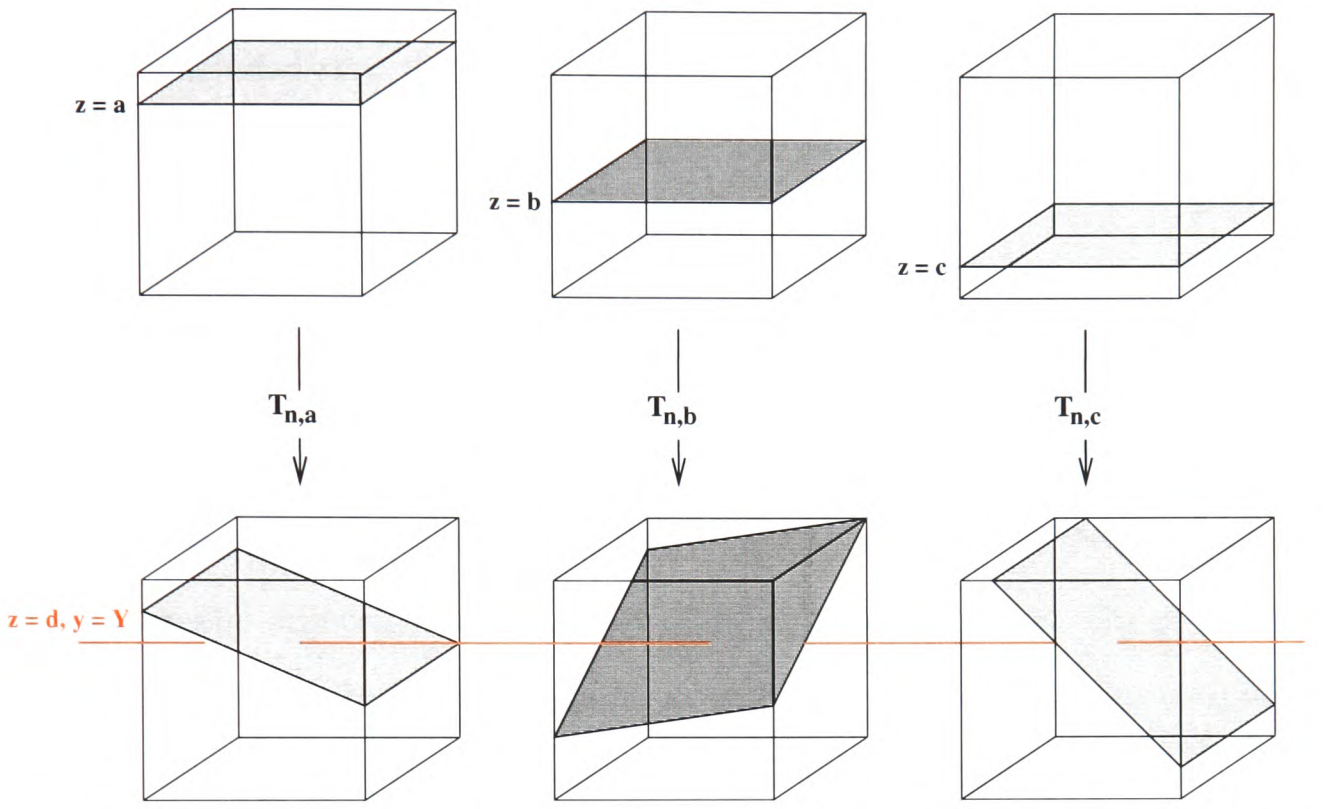


Figure 5.18: Due to the allowed independent movement of slices, it is possible for a single voxel contained within a slice $z = d$ of the realigned data (shown as the intersection of the line $z = d$ with each transformed plane in the bottom row of volumes) to have originated from any slice in the uncorrected data (where $z = a, b, c$ *etc.* in the top row of volumes). It is therefore necessary to check that the matrix $\boldsymbol{T_{n,\sigma}}$, which is used to find the position of the realigned data, transforms the voxel of interest from the slice $z = \sigma$ in the uncorrected data to $z = d$ in the realigned data.

each voxel time-course after the spatial realignment described in sections 5.2.1 to 5.2.4. The underlying acquisition process described in section 5.1 should ensure that once optimisation is complete, all the voxels within each slice of scanner space should contain intensity values acquired at the same time. Unlike existing rigid-body schemes which were shown to be inappropriately used with slice-timing in section 5.1.1, the spatial realignment of the data resulting from the proposed slice-based model allows the application of appropriate temporal interpolation.

Care must be taken, however, to deal with situations where there may be no acquisitions for a given voxel at a particular volume's time-point. In order to apply a shift to reflect the temporal offset of each slice it is necessary to have at

least one sample for each voxel location and time-point. If reconstruction of the data has concluded that the particular voxel was 'missed' during acquisition as a result of subject motion, its intensity value will need to be estimated before the shifted (slice-corrected) value can be calculated.

Time-course Reconstruction

This estimation problem is a natural candidate for spline interpolation methods [113], specifically those which are able to interpolate through their end-points: in this case the adjacent valid samples along each time-course. Given this criterion, a number of choices of spline basis functions are possible. For simplicity, the correction can be performed using Cubic Hermite splines [39] which, given two end-points, describe a curve between and through these points. This allows sample voids to be filled using adjacent values and can be extended to bridge gaps in the realigned time-course where two or more consecutive samples are missing, although in the latter case the potential accuracy will be reduced.

The coordinates $\mathbf{x} = (x, y, z)$ of a Hermite curve for a pair of end-points $\mathbf{x_1} = (x_1, y_1, z_1)$ and $\mathbf{x_2} = (x_2, y_2, z_2)$ can be calculated as:

$$\mathbf{x} = \begin{bmatrix} u^3 & u^2 & u & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \frac{dx_1}{du} & \frac{dy_1}{du} & \frac{dz_1}{du} \\ \frac{dx_2}{du} & \frac{dy_2}{du} & \frac{dz_2}{du} \end{bmatrix} \tag{5.13}$$

where the control parameter u is defined in the range $[0, 1]$. Thus the value of u determines how far $\mathbf{x}$ lies along the curve between $\mathbf{x_1}$ and $\mathbf{x_2}$. This control parameter is set so that the interpolated value along the spline coincides with an

acquisition in the first slice of the volume and the corresponding slice-timing is adjusted to reflect this. Figure 5.19 shows an example of a Hermite spline.

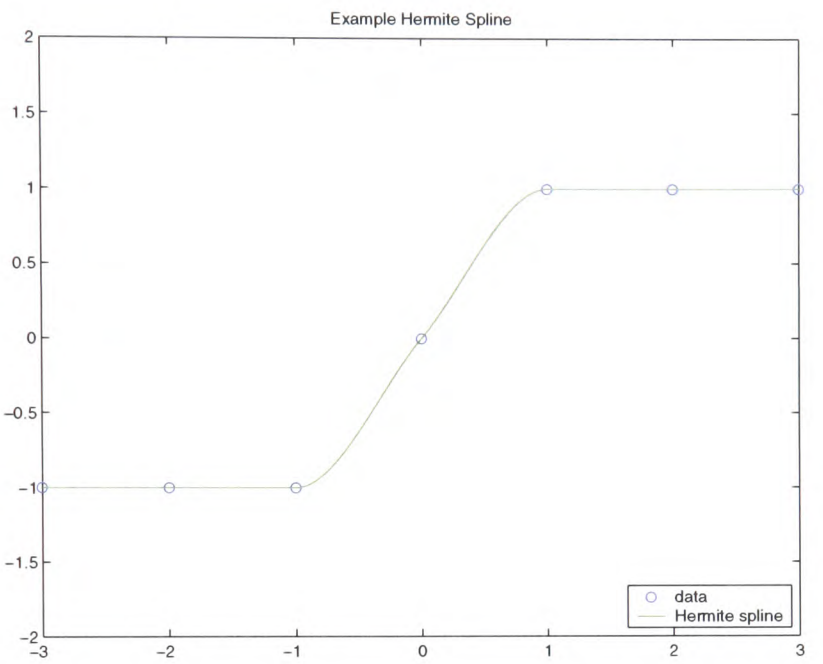


Figure 5.19: The figure shows a Hermite spline interpolating between the original control points (shown as circles). Note that the spline interpolates through the control points, unlike the spline interpolation in figure 4.4.

Temporal Shift

Once the voxel time-courses are complete, the temporal shift associated with each sequentially$y$-acquired slice may be computed using either trilinear or sinc interpolation in 1-D. The erroneous application of temporal offsets when using standard slice-timing corrections, discussed in section 5.1.1, is addressed by maintaining a record of the original z-location of intensity values during the realignment step (section 5.2.3) and it is this value which dictates the amount of temporal shift applied to each voxel, not the z-location of the voxel in the realigned data.

If the voxel value at a location in the time-course has been determined to have been sampled more than once, it is necessary to choose a timing value to associate with this voxel. Theoretically, it is possible to maintain a number of

separate records relating to each contributing location which can be independently interpolated and only then combined. In practice however because of the increased storage and computation requirements, this step is currently omitted.

As with the spatial interpolation, both these temporal interpolation stages are important contributors to the accuracy of the entire spatio-temporal method. Section 6.3.2 presents an accuracy comparison of these two methods.

In the next chapter, the development and testing of a preliminary implementation of the spatio-temporal model are presented. This application of the methods developed above serves not only as a first working spatio-temporal software application, but also as a validation of the techniques which have been introduced in the preceding sections.

Chapter 6

Implementation and Testing

This chapter describes the implementation and testing of a spatio-temporal realignment scheme based on the theory and techniques developed in Chapter 5. The method is first translated into pseudo-code, then each stage is tested, with particular emphasis being placed on the behaviour of the cost function proposed in section 5.2.2. By carrying out these tests using data-sets introduced in section 5.1, which were previously shown to suffer degradation when processed with existing spatial and temporal realignment algorithms, it is possible to demonstrate the clear improvement in accuracy which the new scheme can potentially offer.

The goal accuracy of motion correction in FMRI is typically at the sub-voxel level, that is, of the order of a one or two millimetres in each direction. This is necessary to ensure that consistent voxel-wise comparisons are made across time. Accuracy at an even greater level is often desirable in order to limit the possibility of two or more adjacent voxels, where each corresponds to a different tissue-type or level of activation, being averaged to give a resulting intensity which corresponds to none of the original data.

6.1 Method Summary

The spatio-temporal model for motion correction in FMRI described in section 5.2, now referred to as Temporally-Integrated Geometric EPI Realignment (TIGER), has been implemented in code. The full scheme is described by the pseudo-code in algorithm 2. In order to distinguish between the TIGER model and the implementation being tested here, the latter has been renamed as the FSL Optimised Retrospective Correction Environment (FORCE).

For each volume in the time-series the algorithm begins on line 8 by estimating the gross affine motion in that volume using the MCFLIRT optimisation call described in section 3.1.4 (lines 10 and 12). These estimates are used as the values of T_n which are decomposed using equation 5.1 (line 14). Once this initial slice distribution has been determined, the variance of the remainder of the time-series (excluding the current test volume and the adjacent volume) is determined on lines 16, 17 and 18 which correspond to equations 5.6, 5.7 and 5.5 respectively.

In order to perform this comparison, the overall 4-D variance including the test volume is calculated using equations 5.9, 5.8 and 5.10 and averaged down to a single value on line 25. The optional gradient descent search is represented by lines 27 to 32. At each iteration of the optimisation, described in lines 15 to 33, a perturbation matrix is constructed on line 22 which is combined with the initial affine matrix according to equation 5.2 before decomposition via equation 5.1. These operations are encapsulated by the call on line 24. The optimal value for the perturbation matrix with respect to the cost function is determined by a series of 1D golden section searches over each of the 6 rigid-body parameters in turn. Search in any of the directions halts when a specified parameter on the step size in the 1D search has been reached.

Algorithm 2 FORCE(time-series4d[ ] ts)

```
1: N ← length(ts)
2: S ← z_dim(ts)
3: Matrix[n] T_n
4: Matrix[n*s] T_{n,s}
5: vector(6) params
6: 3dvol SumSqVol, MeanVol, CostVol, TotMean
7: Matrix P ← Identity
8: for i ← 1 to N − 1 do
9:    if i = 1 then
10:      T_n[i] ← opt_trilinear(ts[i], ts[i + 1], 8mm, Identity[4], large_tol)
11:   else
12:      T_n[i] ← opt_trilinear(ts[i], ts[i + 1], 8mm, T_n[i − 1], large_tol)
13:   for j ← 1 to S do
14:      T_{n,s}[S ∗ (i − 1) + j] ← fan_out(T_n[i], j/S)
15: for i ← 1 to N − 1 do
16:   SumSqVol ← sum_sq(ts, i, N, T_{n,s})
17:   MeanVol ← mean_fn(ts, i, N, T_{n,s})
18:   CostVol ← var_cost(i, N, SumSqVol, MeanVol)
19:   for k ← 1 to 6 do
20:      {golden section search}
21:      while opt_tol > limit do
22:         P ← build_perturb(params[k], P)
23:         for j ← 1 to S do
24:            T_{n,s}[S ∗ (i − 1) + j] ← fan_out(P ∗ T_n[i], j/S)
25:         totMean ← cost_upd_opt(ts[i], N, SumSqVol, MeanVol)
26:   for k ← 1 to 6 do
27:      {gradient descent search}
28:      while opt_tol > low_limit do
29:         P ← build_perturb(params[k], P)
30:         for j ← 1 to S do
31:            T_{n,s}[S ∗ (i − 1) + j] ← fan_out(P ∗ T_n[i], j/S)
32:         totMean ← cost_upd_opt_gdt(ts[i], N, SumSqVol, MeanVol)
33:   mat_upd(i, P, T_n, T_{n,s})
34: ts ← correct_voxel_offsets(ts, T_{n,s})
35: for x ← 1 to x_dim(ts) do
36:   for y ← 1 to y_dim(ts) do
37:      for z ← 1 to S do
38:         for i ← 1 to N do
39:            hermite_spline_fill(ts, x, y, z, i)
40: for x ← 1 to x_dim(ts) do
41:   for y ← 1 to y_dim(ts) do
42:      for z ← 1 to S do
43:         interpolate_1d(ts, x, y, z)
```

Once an optimal slice orientation has been found, the individual slice transformations relating to the current test volume are updated with the corresponding perturbation and the adjacent volume's affine matrix is adjusted according to equation 5.3 on line 33. It is then necessary to use the Hermite spline formulation in equation 5.13 to approximate the intensity values at any voxel locations which the scheme has determined were not imaged at all during acquisition (line 39).

Finally the temporal interpolation call on line 43 uses either trilinear or sinc interpolation to shift each voxel value so that the final images reflect an acquisition at one discrete timing per volume. This is a modified version of the conventional slice-timing depicted in figure 5.1 except that the shift applied depends on the location of the slice in scanner space where the data originated from rather than the location determined under the slice transformation.

In the remaining section, the implementation described in algorithm 2 is tested using both the synthetic data described in sections 5.1.3 and 5.1.4.

6.2 Testing

In this section, the ideas introduced to evaluate the accuracy of rigid-body motion correction schemes in section 3.2 are developed to allow comprehensive testing of the assumptions behind slice-based volumetric acquisitions. By choosing suitable test data and evaluation metrics, it is possible to perform a comparison between existing two-stage slice-timing with rigid-body motion correction approaches, described in section 5.1.1, and the novel method which has been proposed in section 5.2.

The evaluation reported below proceeds by evaluating the accuracy achieved by the separate stages of the FORCE implementation. The first of these is the

slice-based decomposition of the original volumetric data, tested in section 6.2.1. The optimisation of this initialised data, implemented according to section 5.2.3, is examined in section 6.2.2 and includes an evaluation of the spatio-temporal interpolation described in section 5.2.4.

Later, a comparison is made between the accuracy provided by trilinear and sinc interpolation methods in the final temporal re-sampling of the data. Finally, given that this interpolation is based on the original z-location (and hence temporal offset) of each voxel within the volumetric data, a demonstration (section 6.3.2) is included to show the incorrect results achieved if this step is omitted.

6.2.1 Initialisation

The slice-based decomposition described in section 5.2.1 is a fundamental part of the spatio-temporal correction and it is possible that much of the spatial realignment necessary will be achieved by this step.

Care must be taken, however, as such an immediate decomposition into separate slices will inevitably tend to bias any further optimisation, and hence the eventual solution, towards the quality of the initialisation itself if the optimisation can guarantee local convergence only. As the pose estimation begins with an estimation of volumetric motion, which is a gross approximation if any between-slice motion has occurred, it is necessary to ensure that the the initialisation is as accurate as possible.

It is possible that gross volumetric estimation may tend to over- or under-estimate the motion contained within a particular volume so the accuracy of initial estimates was tested using the data-set containing nodding motion introduced in section 5.1.4 as well as the in-plane motion data-set from section 5.1.3. In the interests of providing some meaningful comparison with the results from

	In-plane	Through-plane
RMS error /mm (uncorrected)	3.0183	3.2789
RMS error /mm (after init.)	0.4729	0.4804

Figure 6.1: RMS errors in the initial estimation of slice orientation based on artificial data displaying in- or through-plane motion characteristics.

Chapter 3 it is sensible to extend the RMS test measures used previously with the rigid-body motion correction to measure accuracy. Some extensions are required as a spatio-temporal match is now sought, rather than a series of independent spatial registrations.

If a scanner-centric coordinate frame is considered, a range of slice-based transformations has been applied to generate the motion-corrupted data based on original object-centric data (sections 5.1.3 and 5.1.4). The spatio-temporal correction scheme will generate transformations which move the scanner-centric data back to object space. Therefore a measured error between each correction matrix (scanner to object) and the inverse of the corresponding generative slice matrix (object to scanner) of each slice is required. As in the earlier motion correction assessment, this test metric can be used in the case of artificial data where ground truth is known. In practice, the RMS error associated with each slice transformation is summed and averaged over all slices and all time-points.

The estimated slice matrices generated by the stepped-slice initialisation can be compared to the known matrices used to form the motion-corrupted data from sections 5.1.3 and 5.1.4 (figure 6.1) using the RMS measure from section 3.2.1. For the purposes of comparison, the uninitialised data-sets were measured to contain approximately 3 millimetres RMS error due to the induced motion.

Note that because the perturbations, whether rotational or translational, were applied to groups of slices in the high-resolution space according to the subsequent

re-sampling, there is always a discrete matrix which describes the motion of an individual slice in the low-resolution space of the test data-sets.

It can be seen from these interim results that the initial estimation of the intra-volume motion is within less than 0.5 millimetres of the actual value when data with known motion is tested. This is already appreciably close to the known solution (accuracy is of the order of magnitude of the rigid-body corrections reported in figure 3.9) which has the benefit of allowing the subsequent optimisation to focus on fine-tuning the motion parameters at the desired sub-millimetre scale (Chapter 1).

6.2.2 Cost Function Optimisation

It is straightforward to repeat the test measure used in section 6.2.1 to provide a measure of the correction provided by the stages in the spatio-temporal scheme which implement the optimisation of section 5.2.3. This analysis demonstrates the suitability of the cost function developed in section 5.2.2.

In the course of evaluating the FORCE implementation, it is shown that the proposed optimisation scheme suffers from a fundamental degeneracy, specifically that the scheme does not incorporate a canonical reference volume. Having established that it is possible to correctly identify the correct spatio-temporal realignment of the data, an alternative registration scheme is developed which is designed to circumvent this situation. This is described in more detail in section 6.3.

Note that from this point forward, unless otherwise specified, testing was carried out on the 2 degree back-and-forth data-sets so as to keep the motions considered within a typical range for a co-operative subject. Additionally, for all the testing described from this point onward, sinc interpolation was used for

the spatial re-construction of the data under the estimated slice motions. While sinc interpolation involves a higher computational overhead than the simpler choice of trilinear interpolation, it can provide significantly more accurate results (typically to an order of several percent). In practice, it was also found that in situations where the optimisation scheme was already close to the true solution, the increased accuracy obtained by using sinc was able to compensate adequately for the additional computational load when compared with trilinear interpolation which required additional cost function evaluations in order to converge on the true solution.

In order to evaluate the performance of the cost function in detail, the value of the cost function was computed for a range of perturbations around the true solution for each of the six motion parameters. The plots of this cost variation with parameter value are shown, for the first two volumes of the shaking head motion data, in figures 6.2 and 6.3.

For all plots, the cost function is at a minimum near the known realignment of the volume (in this instance, where each perturbation parameter is zero). Discrepancies of a the order of 0.01mm can be seen in a few situations, which can be attributed to the accuracy of the interpolation necessary to compute the cost function on realigned data. In general, interpolation back to the exact original values is impossible, hence it is invariably inaccurate. It is possible to see significantly more discontinuities in the z-translation plot which will be examined more closely in section 6.2.3.

The results of a similar evaluation on the nodding motion data are presented in figures 6.4 and 6.5.

Within the specified ranges, it is not possible to locate the minimum of the cost, in particular the y-translation for the nodding data. This is re-plotted, over

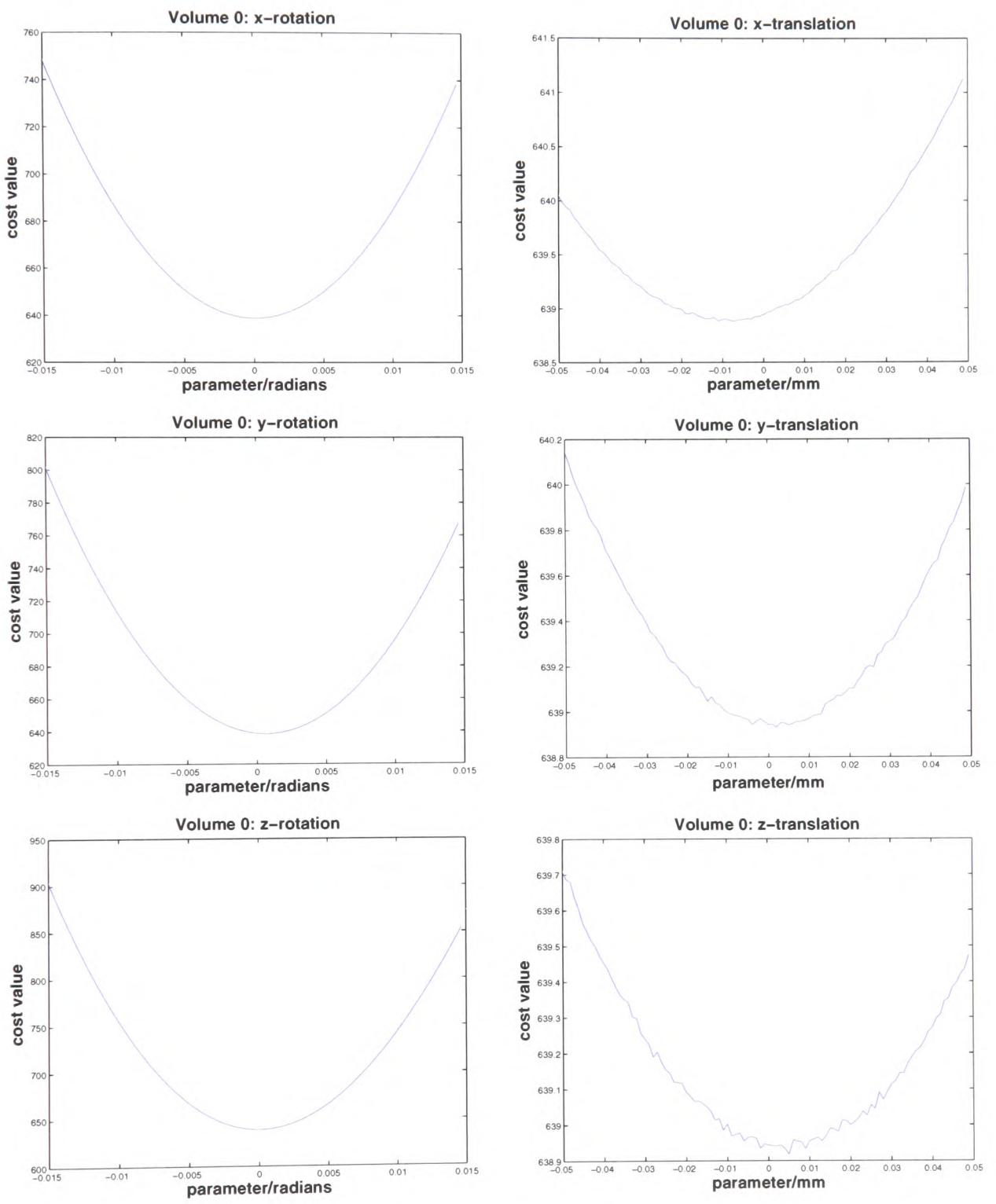


Figure 6.2: Plots show the variation of cost function value with each of the six motion parameters in turn as applied to the first volume of the shaking motion data-set. Note the slightly offset minima in the case of y-rotation and the translations. Also note the additional discontinuities in the z-translation due to the lower sampling resolution in this direction in the original data.

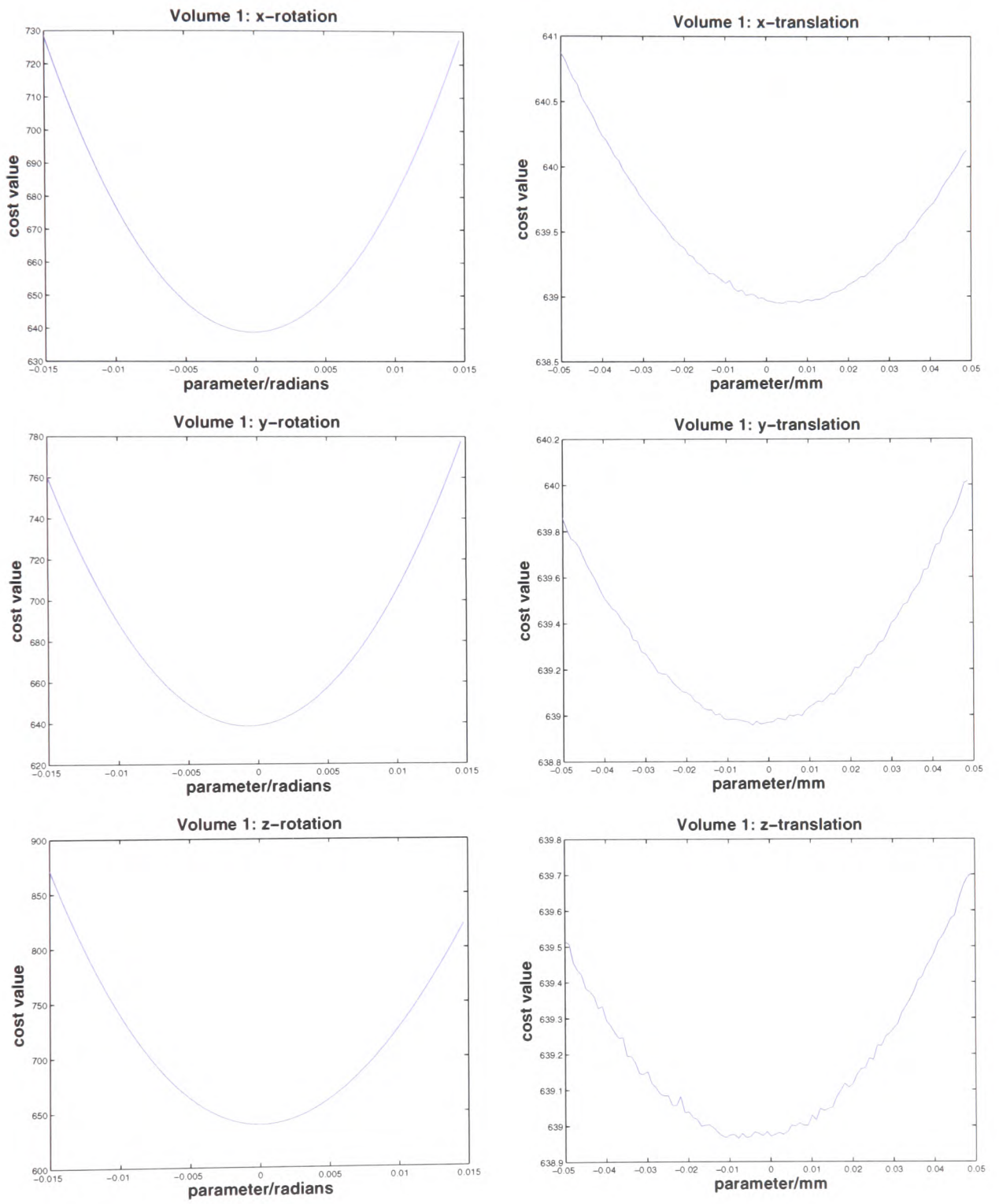


Figure 6.3: Plots show the variation of cost function value with each of the six motion parameters in turn as applied to the second volume of the shaking motion data-set. Note, as with the previous figure, the slightly offset minima in the case of y-rotation and the translations.

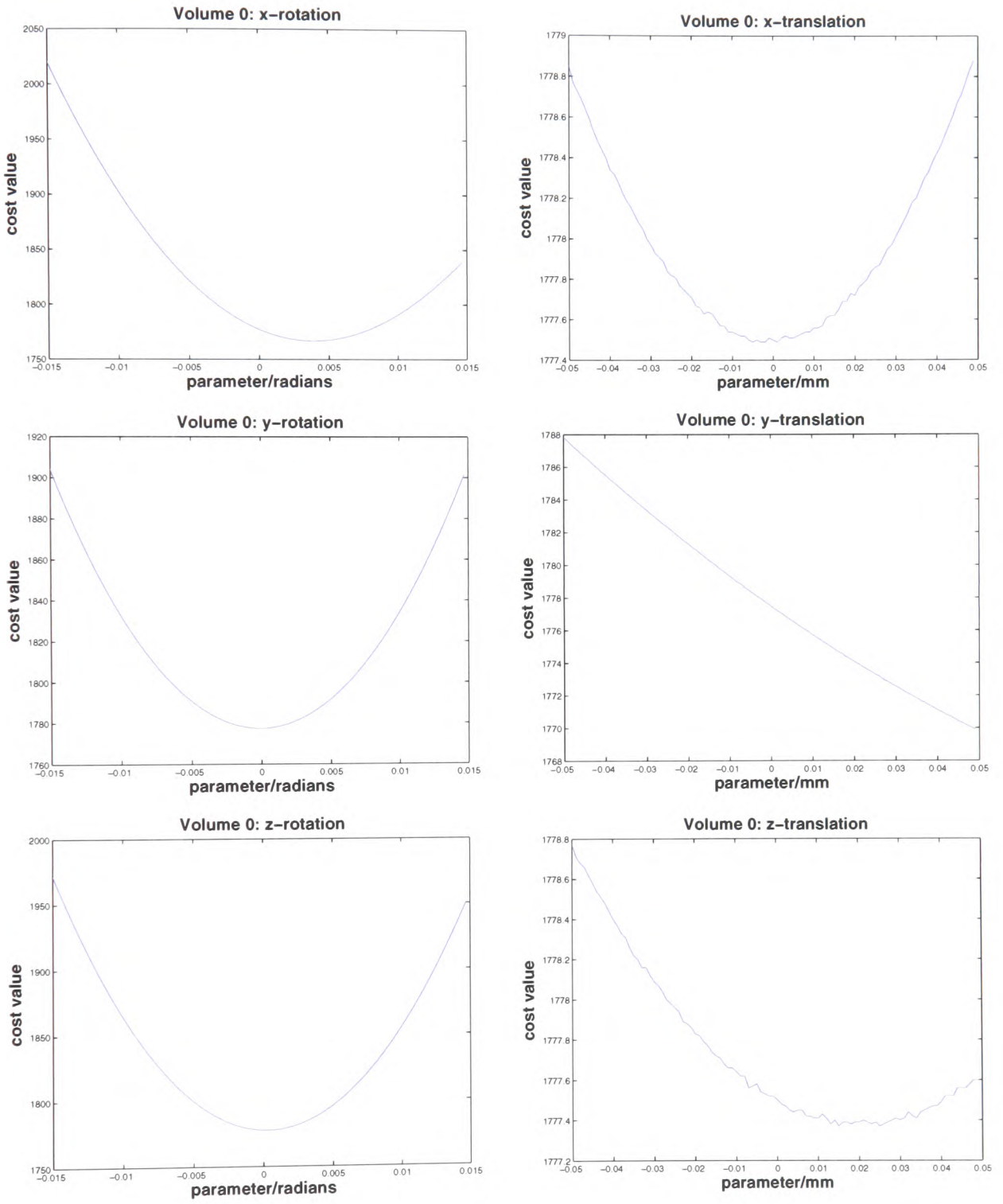


Figure 6.4: Plots show the variation of cost function value with each of the six motion parameters in turn as applied to the first volume of the nodding motion data-set. The obvious difference in these plots compared to the previous figures for the shaking data is the large error in y-translation likely to arise from a shortening of the data in this direction due to the nature of the motion present.

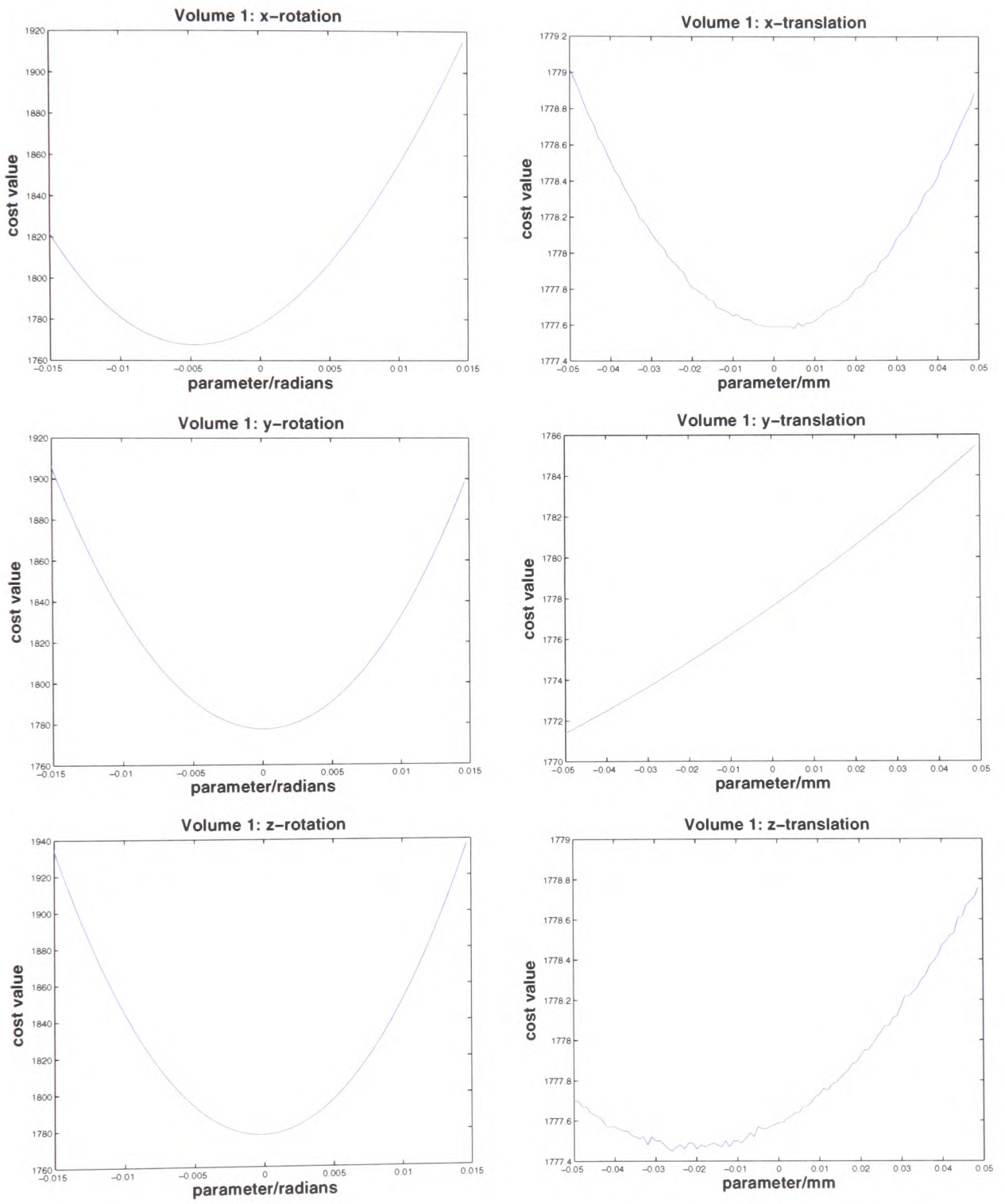


Figure 6.5: Plots show the variation of cost function value with each of the six motion parameters in turn as applied to the second volume of the nodding motion data-set. The errors in the position of the minima are similar to those in figure 6.4, with the exception of a change of sign due to the symmetry of the two motions.

a larger range of perturbations, in figure 6.6.

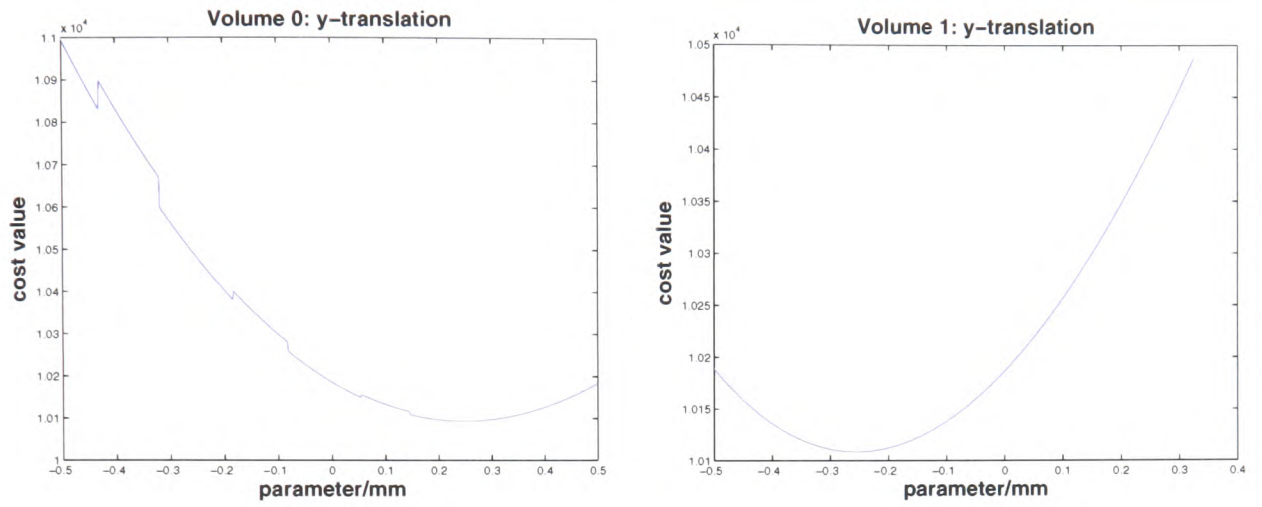


Figure 6.6: Plots show the variation of cost function value with y-translation over a larger scale than before as applied to the first and second volumes of the nodding motion data-set. This re-plotting allows the position of the cost function minima, not visible at the original plotting scale, to be identified.

One explanation for the offset minimum in the case of the nodding motion is that, in the course of a forward pitch, the y-coordinate of voxels within a slice will shift (figure 6.7). As the true solution is known to involve such a motion, which requires interpolation along this direction, this may be a potential source of error. If this were the case, one would also expect a similar shift in z-translation, which can be seen from the plots in figures 6.4 and 6.5. It is assumed that the difference in the size of offset associated with these two parameter results from the discrepancy in sampling frequencies. The coarser resolution in the z-direction will lead to smoother estimates while the finer y-resolution will tend to be more influenced by interpolation errors.

6.2.3 End-slice Interpolation

It can be observed that the plots of cost function value with varying z-translation exhibit a noticeably less smooth profile than the plots generated by perturbing

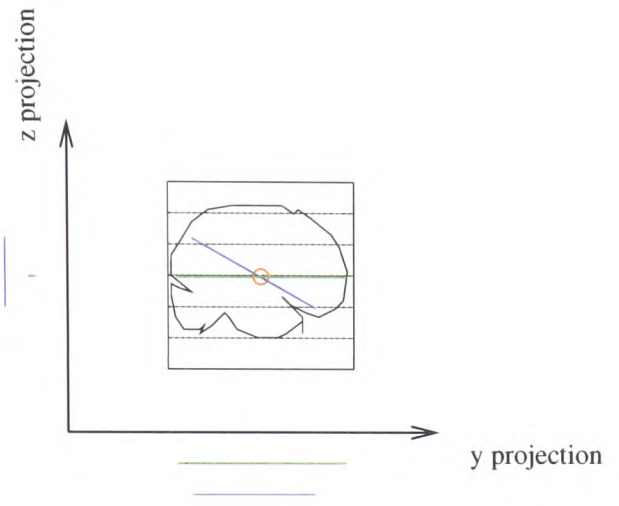


Figure 6.7: In the case of a pitching movement, equivalently a rotation about the x-axis projected through the centre of gravity of the head (approximately shown as a red circle), the y-coordinate and z-coordinate of voxels within that slice will shift.

the other 5 motion parameters. This is due to the lower sampling resolution in the z-direction. Earlier work in section 3.1.5 showed how movement out of the FOV of the images could lead to potentially more inaccurate interpolated values after realignment. By extending the idea of padding the data with duplicated end slices to the spatio-temporal scheme, it should be possible to show significantly smoother cost function behaviour (which will, in turn, lead to a simpler optimisation space) in figure 6.8.

The bottom left plot in figure 6.8 seems highly abnormal on first examination: it is likely that the nodding motion (which sweeps through slices in the z-direction) is interacting with the additional end slices. This interaction manifest itself as a set of stepped cost values (changing as the number of slices which rely on the padding slice increases or decreases) where the calculation in these slices would otherwise be zero.

It is obvious that the plot shown for the z-translation of the second volume of the shaking data (top right in figure 6.8) is not exhibiting a minimum at the desired location. To investigate this further, the variation of cost with z-

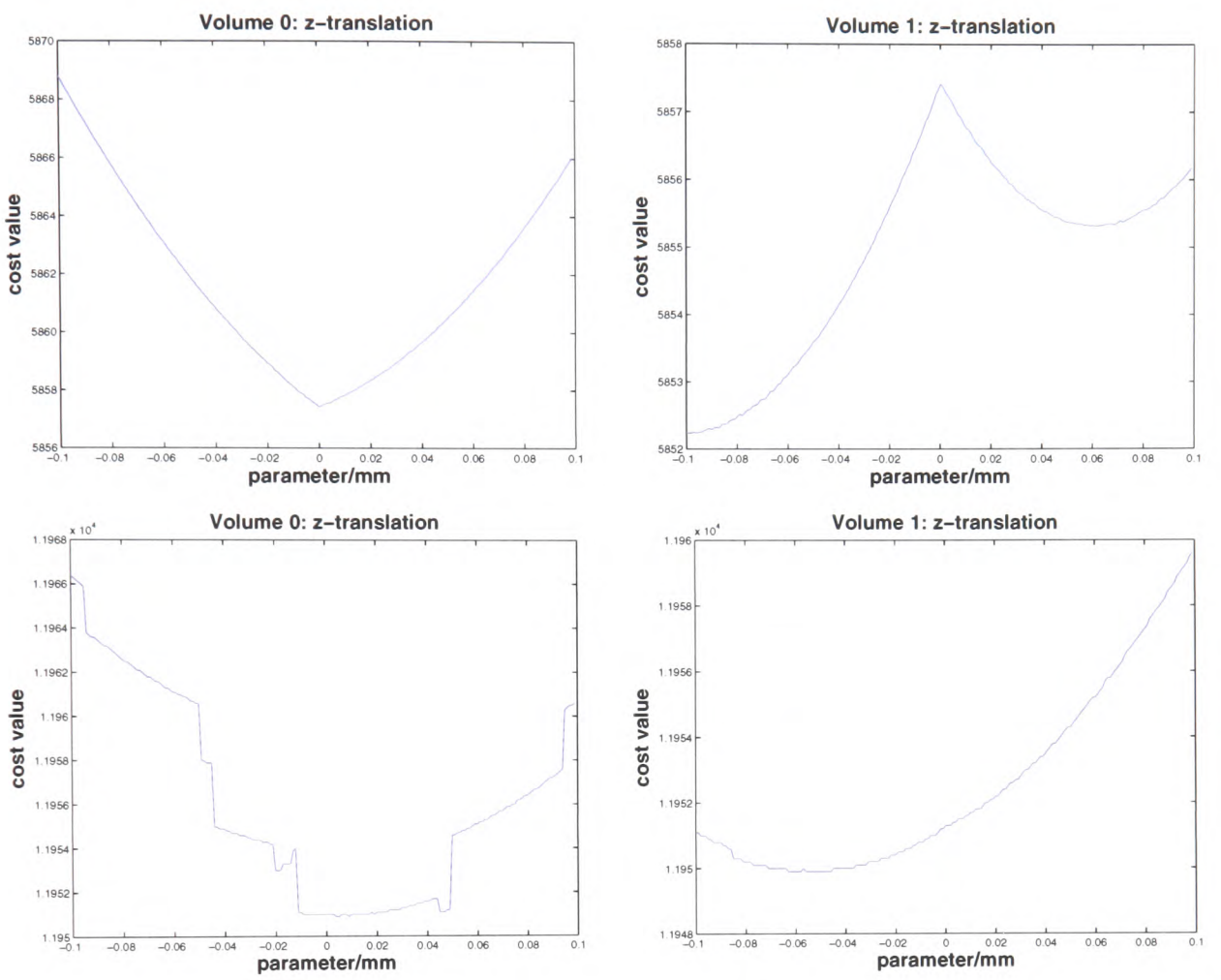


Figure 6.8: Plots show the variation of cost function value with z-translation after the inclusion of end-slice padding. Plots on the top row correspond to the first two volumes of the shaking motion data while those in the bottom row correspond to the first two volumes of the nodding data. Note the reverse peak in the top right-hand plot, thought to be caused by an overlap of sample grids, and the highly discontinuous plot at bottom left caused by the interaction between the through-plane motion and the additional slices.

translation was re-plotted, over a larger range of perturbations, in figure 6.9.

It is now apparent that for the second volume of this data set, the cost function decreases towards the minimum and then suddenly peaks. This unusual behaviour has been observed in earlier work on medical image registration using mutual information by Pluim *et al.* [119]. It was shown in this work that a registration measure can show sudden changes of value for transformations which precisely align the sample grids of the two images. These patterns of local extrema impede the registration optimisation process and more importantly may

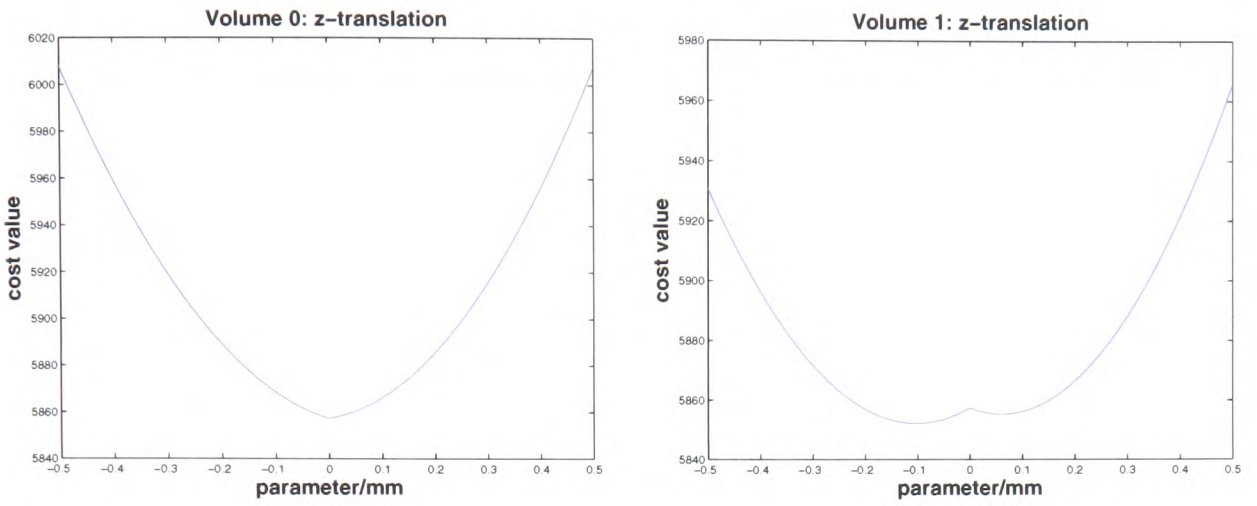


Figure 6.9: Plots show the variation of cost function value with z-translation after the inclusion of end slice padding. Plots correspond to the first two volumes of the shaking motion data, computed over a larger range than before. While the plot on the left is well-behaved at a larger scale than shown in figure 6.8, the plot on the right now displays an erroneous peak where a minimum is expected. The increased range has allowed this to be identified as a local rather than a global peak, a characteristic not immediately obvious at the earlier plotting scale.

rule out sub-voxel accuracy. One explanation for these observations is that by interpolating between grid locations, noise and other small structures, which would otherwise penalise the cost function, are often smoothed out. A simple approach to avoiding the situation of grid alignment is to re-sample one of the images so that grid-aligning transformations cannot occur. The authors of this research also noted that although such a re-scaling will dramatically improve the smoothness of the cost function, it will not necessarily lead to improved registration results. Given the already low resolution of FMRI data, it is unclear if such resamplings are practical and while the inclusion of end slice padding is able to improve the behaviour of certain costs (for example, the z-translation behaviour in the first volume of the nodding data set), it may also lead to a less smooth optimisation profile, as seen in figure 6.9. While the latter observation holds, it is sensible to retain a conservative approach by omitting the padding option. In this case it is necessary to re-plot the cost behaviour with z-translation for the second volume

of the nodding data-set in order to confirm that it has a minimum close to the case of no perturbation. This is shown in figure 6.10.

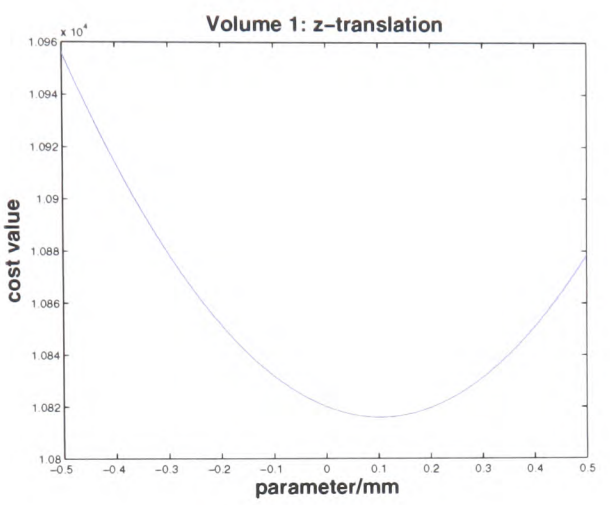


Figure 6.10: Plots show the variation of cost function value with z-translation, over a larger range of perturbations than before, as applied to the second volume of the nodding motion data-set. This figure confirms that the function is well-behaved without the end-slice padding which was evaluated in the previous figure.

6.3 Scheme Degeneracy

In the course of evaluating this scheme, it became apparent that without any additional information (specifically a canonical reference volume in order to enforce an anatomical realignment) the registration scheme suffers a fundamental degeneracy: specifically, that any number of consistent realignments across the time-series will satisfy the minimum variance condition imposed by the cost function. In the case of the data for which the scheme is designed, it is entirely possible that without a very accurate initialisation, the optimisation scheme may drive the registration to a solution which, while consistent in time, is not registered within individual volumes and which is anatomically implausible.

A second issue is that of the reference data used at each stage of the re-

gistration. In the present implementation, each volume is co-registered to the remainder of the time-series (excluding that volume). This relies on the reference volumes being very close to their final position if the multi-dimensional search is to reach a global optimum for the target volume under consideration. Clearly this situation is unlikely, given that even if in some cases the initialisation examined in section 6.2.1 is able to halve the residual error in the data, the initialised volumes will not be close enough to the global minimum to guarantee a robust registration.

Furthermore, there is a range of inter-slice motions for which a rigid-body volumetric estimate cannot realistically be expected to be able to represent, even approximately. Such motions include those tested previously and indeed similar initialisation tests on motion exhibiting the same type of motions but over different ranges often failed to improve on the uncorrected error. More details are presented in the next section, where concerns over the accuracy of an initialisation based on volumetric MCFLIRT estimates are addressed. The issue of initialisation may be tractable if a more exhaustive optimisation search strategy, such as simulated annealing, is implemented. The potential advantages of such a strategy are discussed in greater detail in section 7.1.2.

The issue of optimisation aside, section 6.3.1 addresses the issue of degeneracy which has been raised above.

6.3.1 Static Initialisation

In order to enforce the anatomical constraint, that is, that the realigned data should represent the data in object space, not simply in a pose which satisfies the cost function, it is necessary to include at least one volume which contains little or no inter-slice movement. While this may seem to be a very specific case

of the likely input data, it is relatively easy to select such a volume from within the time-series by examining the relative RMS difference between each volume, as estimated by MCFLIRT. In the case where the relative movement between two volumes is very low, an assumption can be made that little or no inter-slice movement occurred over the course of the acquisition of either. Thus the images will hopefully contain a physically-consistent representation of the brain, even if its global pose is not entirely co-incident with the coordinate axes.

Either of these two internally static images can then be used as a canonical reference volume for the spatio-temporal realignment. An appropriate modification to the realignment scheme is to select the static volume using the procedure described above and to then allow the adjacent volume to perturb according the the model shown in figure 5.17. Once this second volume has been realigned with respect to the first static volume it is possible to use both these volumes as a reference image in the variance cost for the realignment of the third volume. In this way, the reference image is built up iteratively but anatomical constraints are enforced by virtue of starting with a volume which has been measured as the most likely to contain little in the way of inter-slice movement.

A scheme which follows such an approach was implemented. A further modification to the scheme proposed in section 6.1 was to permute the order of the three rotation directions. In this way it is possible to start the optimisation search for a given volume in all six possible permutations of the three rotation parameters. This should ensure that a rotation which does not correspond to the true motion, but which is searched first, is not (incorrectly) selected and then adjusted to a local minimum by subsequent translation searches. Again, evaluation of the original implementation of section 6.1 revealed that this situation arose on a number of occasions when realigning the test data, often resulting in the scheme

being driven away from the global minimum.

In order to test the new scheme, the earlier shaking data-set was modified with the addition of an unperturbed volume at the start of the time-series. To confirm that the cost function would still be able to identify the correct realignment, each parameter was allowed to vary about the known minimum as before. The results of this evaluation are plotted for the second and third volumes of the new shaking data set in figures 6.11 and 6.13 respectively.

As it is not possible to see the minimum in the case of the x-translation for the second volume, this cost variation was plotted over a larger scale of translations and is shown in figure 6.12. In this case the measured minimum is 0.259 millimetres from the correct position. Once again, it seems plausible to attribute these deviations to the inaccuracy of the interpolation method when applied to such a sparse data-set.

It is also necessary to evaluate the non-local behaviour of the cost function to confirm that the scheme will still converge, even if initialised a significant distance from the true minimum. Plots of parameter variation for rotations of $+/-3$ degrees and $+/-2$ millimetres (approximately 1 voxel dimension) are presented in figures 6.14 and 6.15.

The final optimisation test was done to ensure that the multi-dimensional search, described in section 5.2.3 above, was able to locate the minimum cost plotted in the preceding figures. For the shaking data-set containing the canonical reference volume, the scheme was able to identify accurately the dominant direction of movement (in this case, the rotation about the z-axis) regardless of the order of parameter searches. It is likely that the permutation search will have a more significant bearing on realignments of volumes containing combinations of multi-axis rotations and/or translations.

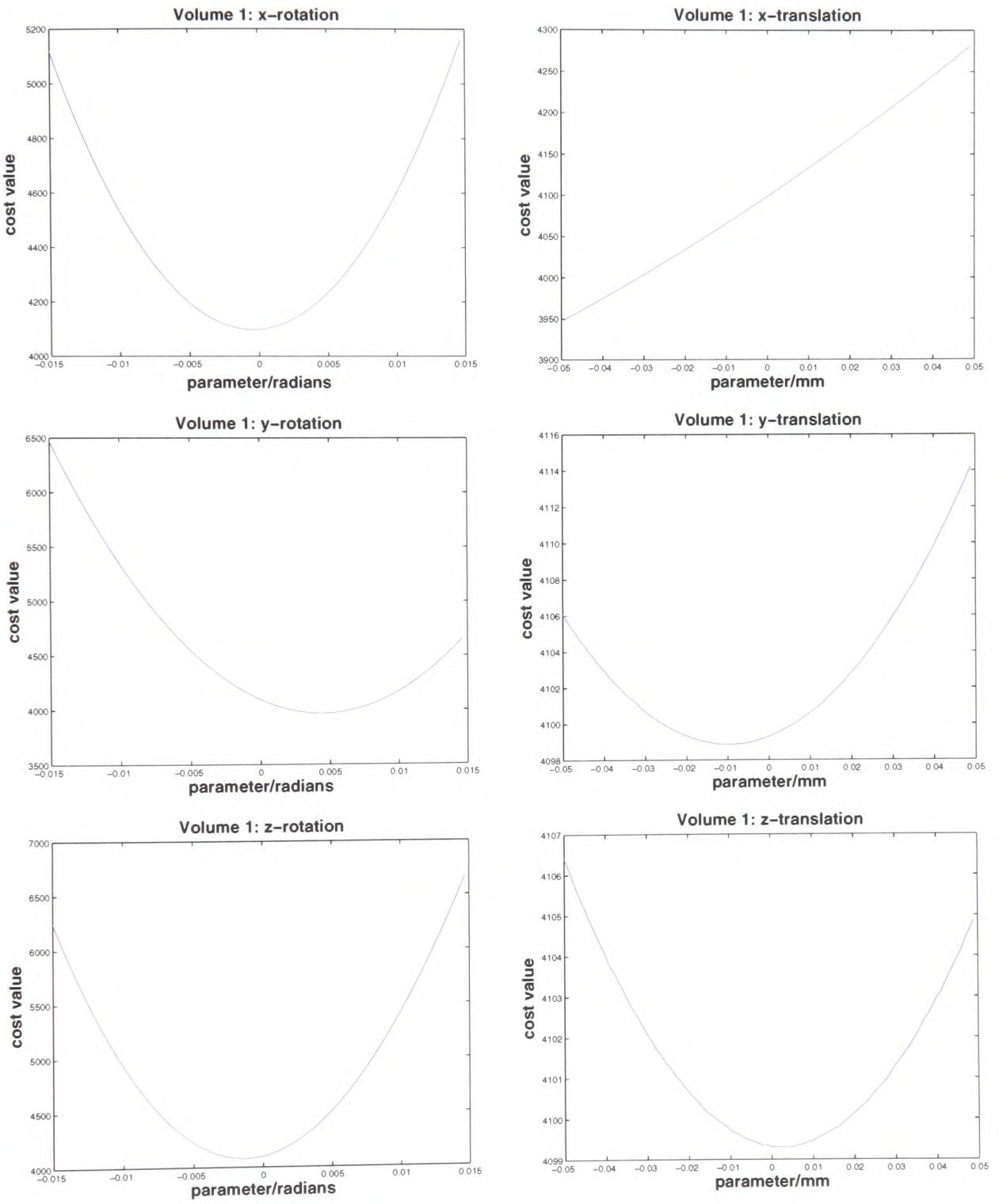


Figure 6.11: Plots show the variation of cost function value with each of the six motion parameters in turn as applied to the second volume of the shaking motion data-set including a static first volume. Note that the behaviour is generally smooth, with small errors when compared to the true cost function minima, but that the minimum for x-translation is off-scale, requiring re-plotting in figure 6.12.

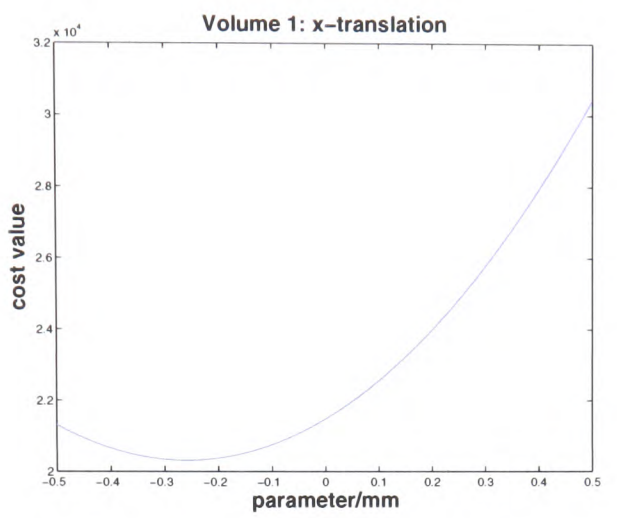


Figure 6.12: Plots show the variation of cost function value with x-translation, as applied to the second volume of the shaking motion data-set including a static first volume. This plot confirms that the cost reaches a minimum approximately 0.25 mm from the true value.

The results of this test revealed that the realignment scheme was able to reduce the average RMS error within each volume from an uncorrected value of 1.9331 millimetres to 0.2829 millimetres. The deviations from ground truth in the parameter estimates were noted to consistently follow the variation in minima shown in figures 6.11 and 6.13. The MARE measure also fell from 16.69 before correction to 11.63 afterwards.

Qualitative examination of the realigned data revealed that very little perceptible global motion remained within each slice. Of potentially greater interest was the presence of local detail variation between time-points. This reflects the changing role played by interpolation across a range of slice poses and further reinforces the earlier hypothesis that interpolation is responsible for much of the error with respect to zero motion observed in the cost plots. The issue of interpolation will be examined more closely in Chapter 7.

The procedure described at the beginning of this section was repeated with the nodding data-set by appending a single static volume to the start of the

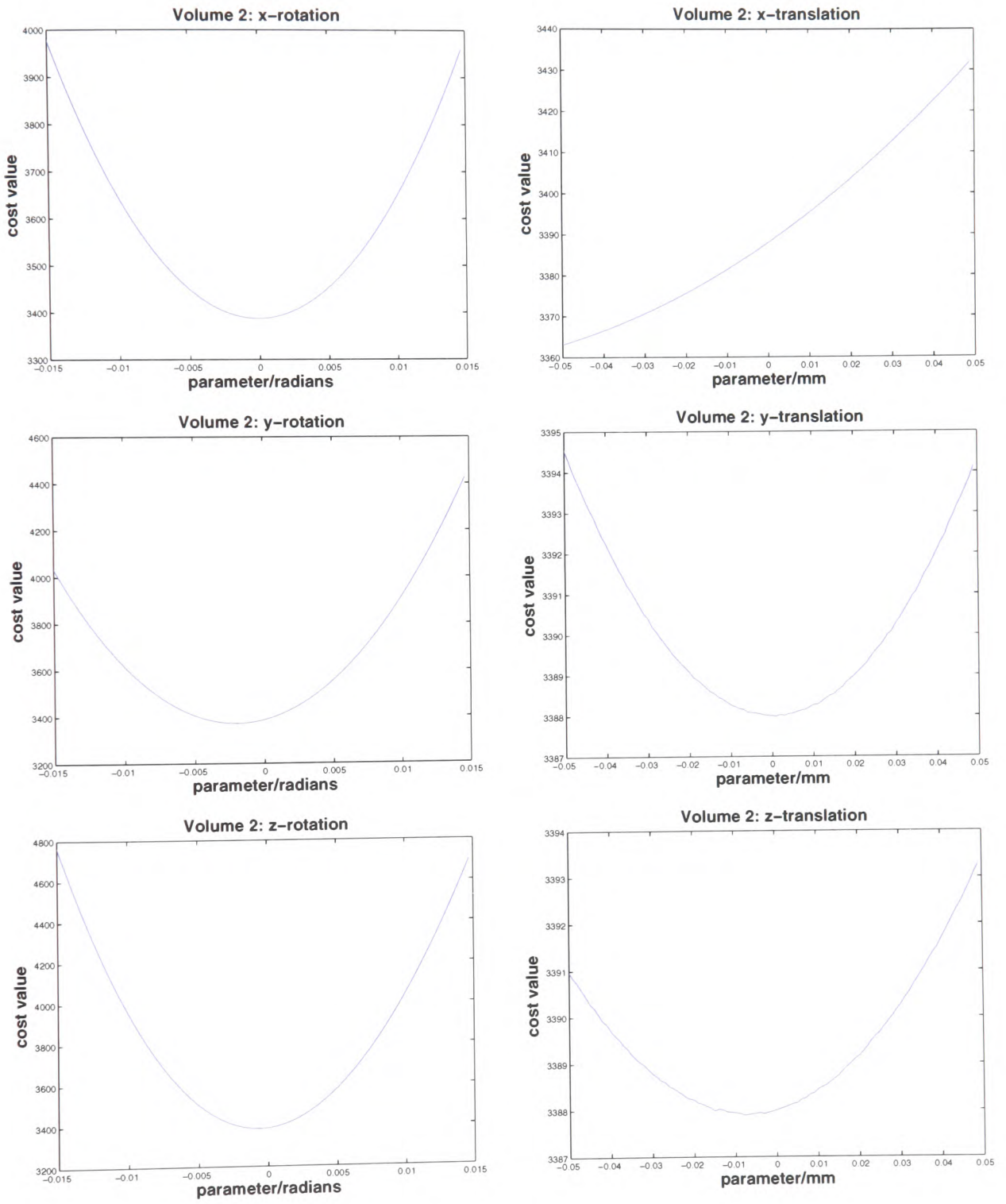


Figure 6.13: Plots show the variation of cost function value with each of the six motion parameters in turn as applied to the third volume of the shaking motion data-set including a static first volume.

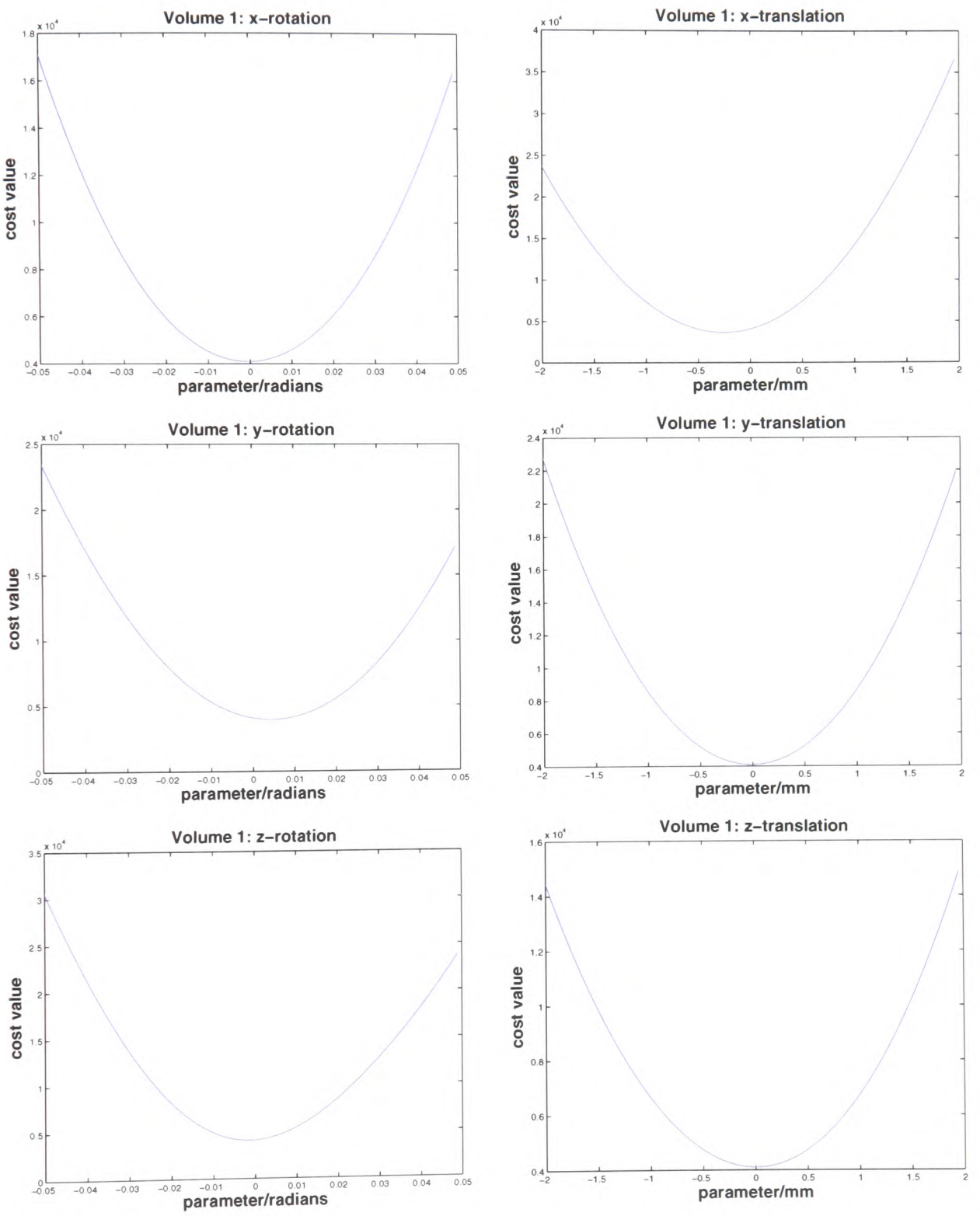


Figure 6.14: Plots show the variation of cost function value with each of the six motion parameters in turn (over a whole-voxel range) as applied to the second volume of the shaking motion data-set including a static first volume.

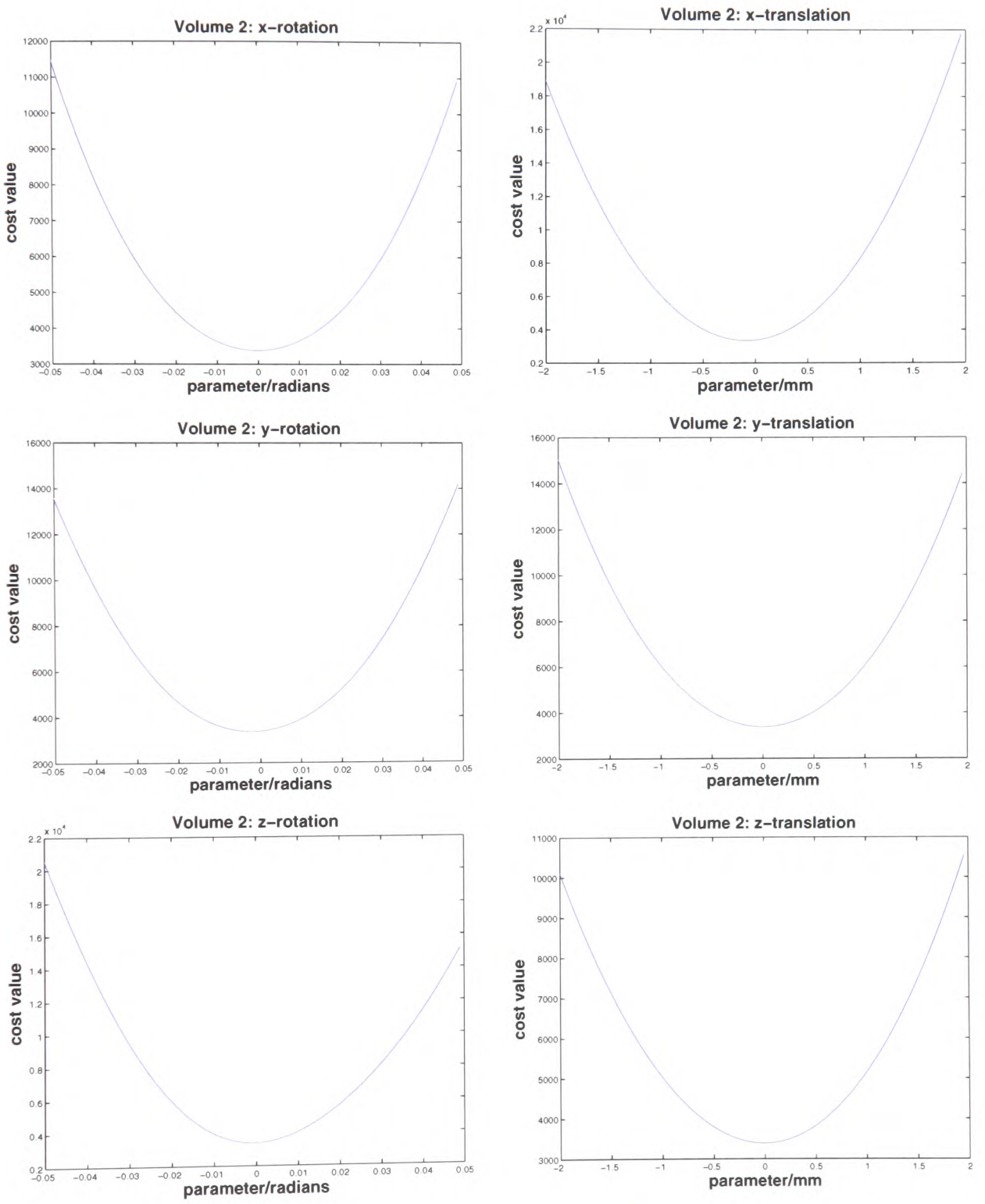


Figure 6.15: Plots show the variation of cost function value with each of the six motion parameters in turn (over a whole-voxel range) as applied to the third volume of the shaking motion data-set including a static first volume.

time-series. Local and more global cost behaviour can be seen again in figures 6.16, 6.17, 6.18 and 6.19. In general, the behaviour of the cost function with rotation parameter variation is both smooth and symmetric about the true minimum. It is still possible to observe off-sets in the minima of translation parameter plots, specifically for y- and z-translation. These are the directions of greatest translational movement associated with a rotation about the x-axis (as is the case for this nodding data) which supports the argument presented in figure 6.7 regarding the limitations of interpolation on low-resolution data, such as the EPI images under examination.

While deviations from zero, consistent with figures 6.16 and 6.17, were still observable in the parameter estimates, the realignment was able to reduce the average RMS error from 2.1863 millimetres to 0.7451 millimetres. Again it was possible to see a reduction in the MARE measure which fell from 18.26 to 14.08 after realignment.

6.3.2 Spline Accuracy

There are two distinct stages to the temporal interpolation of the data after the spatial realignment of sections 6.2.1 and 6.2.2: the spline-based filtering of each time-course, described in section 5.2.5, and the temporal slice-timing where each voxel is shifted along its time-course in an amount proportional to its slice's time of acquisition within the volume. In this section, the accuracy offered by the spline interpolation is investigated, as the test data currently do not contain any temporal variation when realigned correctly.

In order to test the accuracy of the spline interpolation, it is necessary to compare the MARE measure for the data before realignment, after spatial realignment but before spline reconstruction, and finally again after the spline reconstruction.

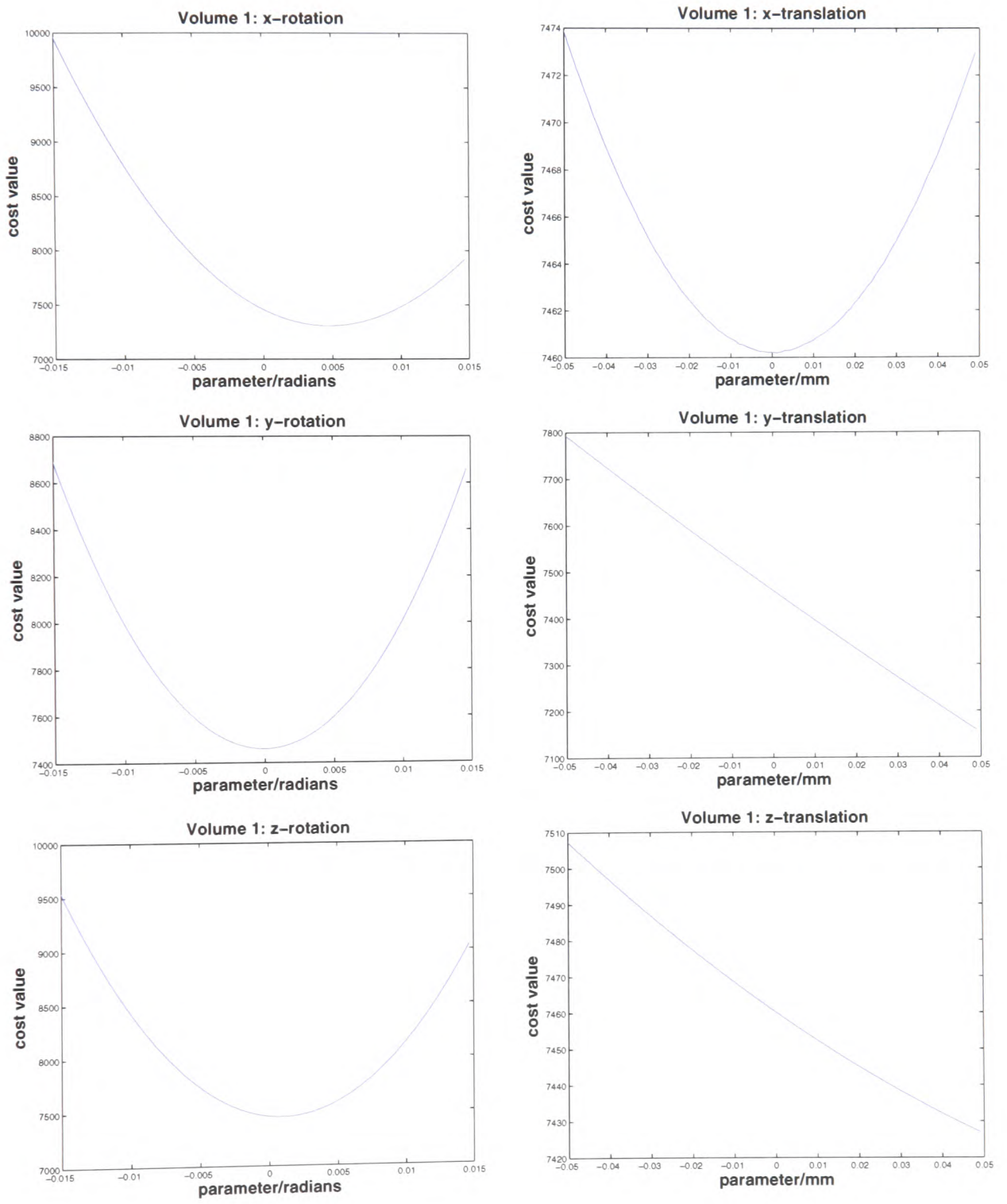


Figure 6.16: Plots show the variation of cost function value with each of the six motion parameters in turn as applied to the second volume of the nodding motion data-set including a static first volume.

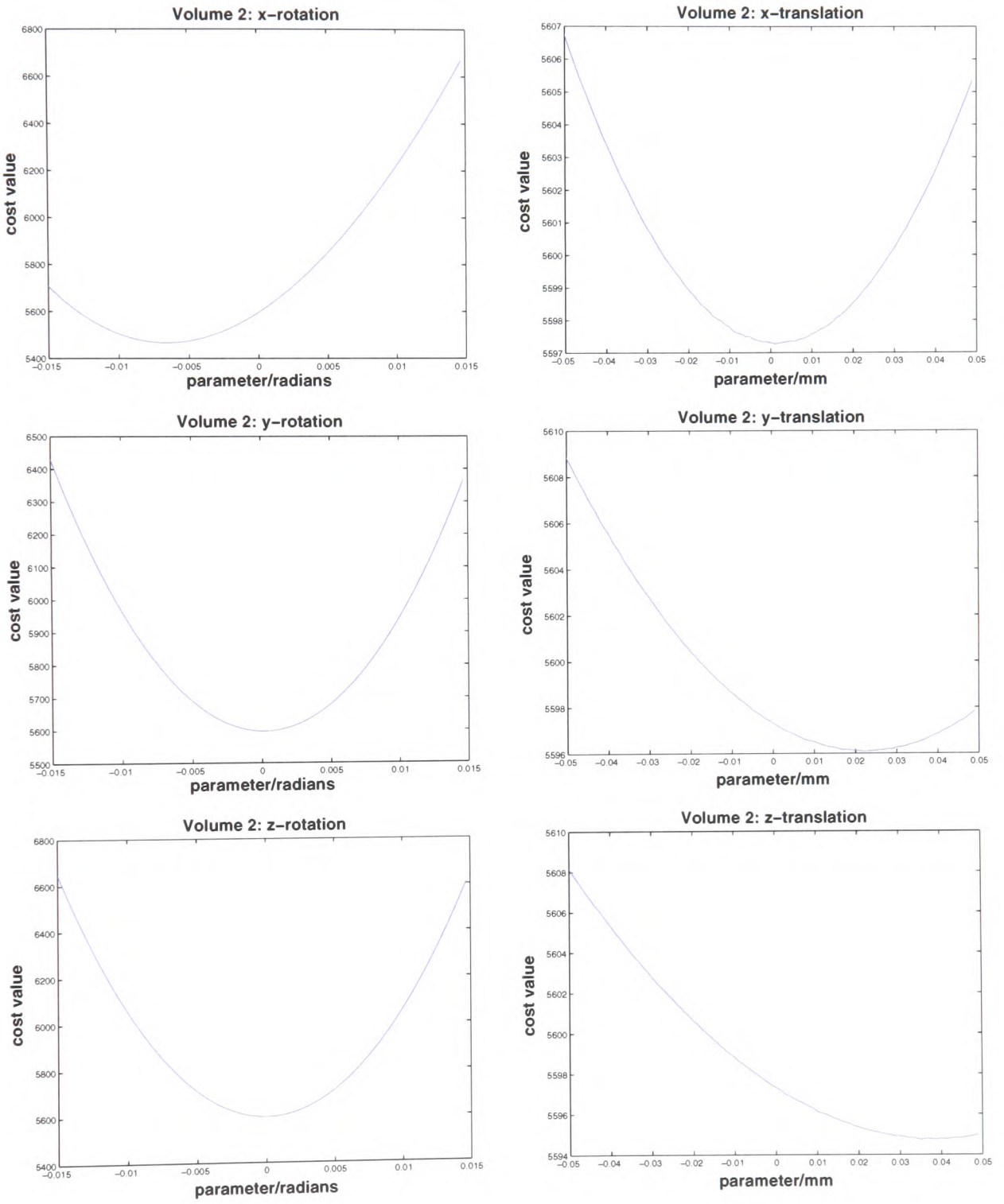


Figure 6.17: Plots show the variation of cost function value with each of the six motion parameters in turn as applied to the third volume of the nodding motion data-set including a static first volume.

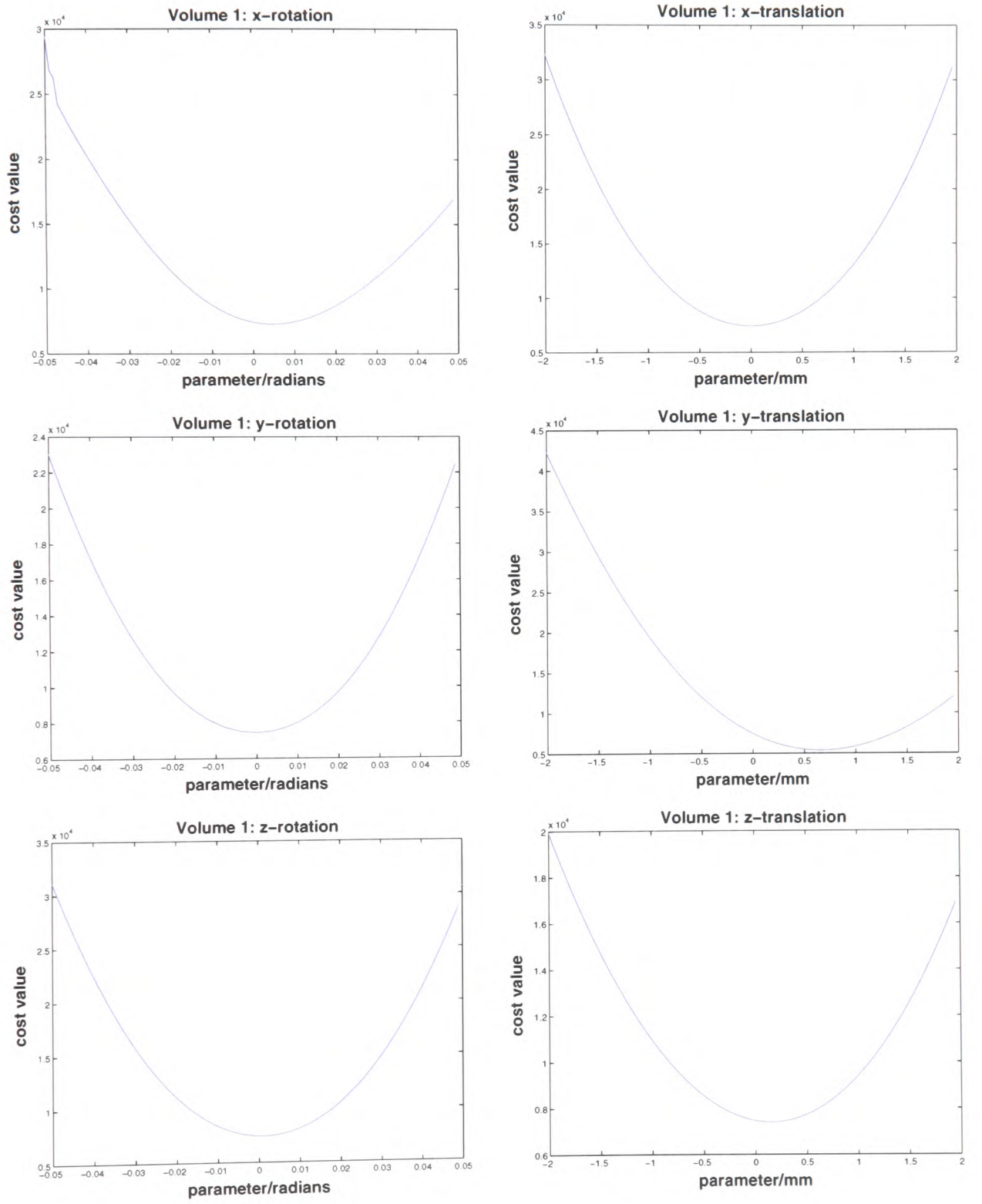


Figure 6.18: Plots show the variation of cost function value with each of the six motion parameters in turn (over a whole-voxel range) as applied to the second volume of the nodding motion data-set including a static first volume.

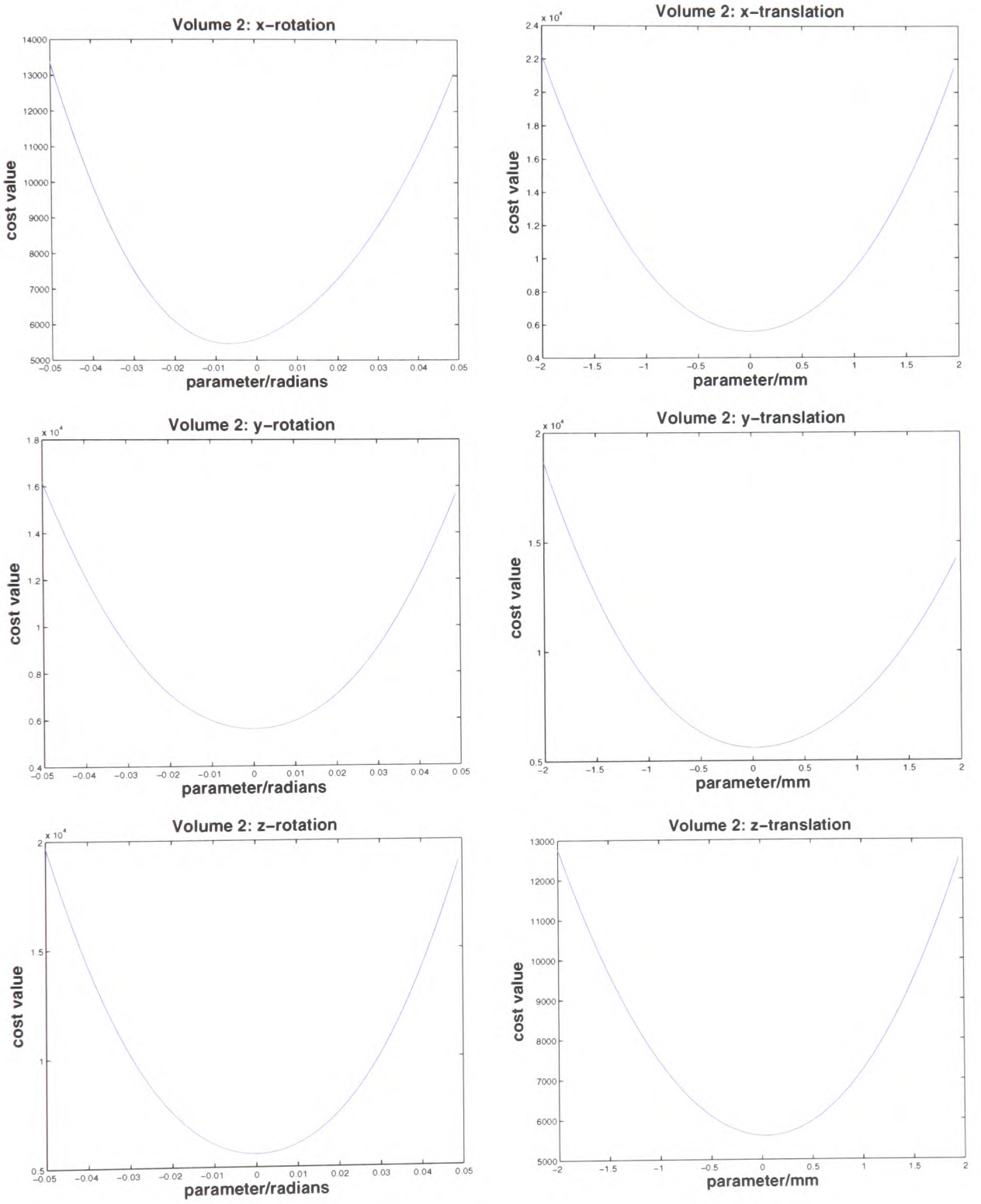


Figure 6.19: Plots show the variation of cost function value with each of the six motion parameters in turn (over a whole-voxel range) as applied to the third volume of the nodding motion data-set including a static first volume.

	Uncorrected	Post-realignment	With Splines
Shaking Data	14.2082	7.2197	7.1796
Nodding Data	16.0322	12.3656	12.2599

Figure 6.20: The table shows the decreased residual error (measured as MARE) which results from time-course correction using splines when compared to spatial realignment alone

For the original nodding and shaking data-sets of sections 5.1.3 and 5.1.4, the results are given in figure 6.20.

6.4 Overall Performance

It has been shown that an implementation of the TIGER model, referred to as FORCE, can be applied in order to provide corrections which are known to be fundamentally more accurate than existing rigid-body and separate slice-timing methods. In the case of both nodding and shaking motion, the adapted static inititialisation scheme (section 6.3.1) was able to significantly reduce the levels of motion in the data to a sub-millimetre level which is necessary for accurate activation detection.

Given that the current implementation is currently untested for the general case of unknown multi-axis, inter-slice motion, a number of ways in which the model may be implemented for use on existing and future clinical data are proposed in Chapter 7. The scheme is sensitive to initialisation and has accuracy limitations due to interpolation, but better optimisation and other measures have the potential to make it usable in real FMRI studies.

Chapter 7

Further Areas of Investigation

This chapter begins with a review of the spatio-temporal scheme, implemented in Chapter 6, and suggests a number of areas in which it can be developed to provide potentially more accurate and robust results on FMRI data. While results have been promising, and indeed form the first attempt to integrate the related problems of retrospective slice-timing correction and spatial realignment, there is still considerable work to do before the scheme can be applied to real functional data-sets, more specifically clinical data, with the required degree of accuracy and desired improvement in activation detection. The following sections seek to identify areas of weakness in the current implementation and draw together the salient points which should lead to an improved algorithm based on the TIGER model and more useful results.

Later in this chapter, a number of possible extensions both to the model and its current implementation are discussed with a view to motivating further research into this important area. By offering a model-based correction for spatio-temporal acquisition and motion artefacts, the proposed framework opens the way to an integrated approach to correcting a number of other sources of artefact. These sources, while usually of secondary concern to the greater confounds caused by

motion, now play a relatively more significant part in the composition of motion corrected data and merge with existing techniques, or in some cases, motivate avenues of further research.

7.1 Extensions to Current Method

In Chapter 5 a new model for motion in FMRI was developed, based on observations regarding the fundamental errors in applying rigid-body motion correction and separate slice-timing interpolation. An implementation based on this model (TIGER) was developed and it was shown, in Chapter 6, that certain key elements of the scheme were able to perform accurately on data exhibiting the general case of inter-slice motion.

It has not been possible, as yet, to provide an implementation which can operate automatically on a wide range of data and in the following section a number of ways forward are proposed with the intention of motivating the development of a working tool which can be used to provide accurate and fundamentally correct realignment to FMRI data. The topics are presented in approximate order of priority, that is, the later topics should not be prerequisite to the provision of accurate results but they may offer some increased flexibility in an implementation based on the TIGER model.

7.1.1 Spatial Interpolation

In the same way that the accuracy of the MCFLIRT scheme improved when sinc interpolation was used relative to the more standard trilinear interpolation, an improvement in the spatial estimation of the TIGER framework (section 5.2.4) is likely to result from the inclusion of a more sophisticated interpolation tech-

nique. The issue of accuracy of interpolation has been highlighted repeatedly as a fundamental factor in the reliability of motion correction, even more so than in conventional medical image registration [73]. Advanced forms of spatial interpolation which have been developed include the estimation of high spatial frequency components within a Bayesian belief propagation framework [40] and the use of Markov [11] and Gibbs [62] random fields for spatial reconstruction. It is plausible that these methods could be adapted to include the robust estimation of inhomogeneous sampling described in section 5.1.2 rather than having to use a separate spline interpolation stage.

An alternative approach to the problems caused by inaccurate interpolation might alternatively involve deriving the interpolated values from the original k-space data. By interpolating in k-space, before the data is Fourier transformed into cartesian image space, it might be possible to achieve more accurate interpolation by avoiding the errors associated with interpolating truncated data in the form of images.

7.1.2 Alternative Optimisation Framework

It can be seen from the cost function plots in Chapter 6 that the cost function contains some local minima which may confound the simple parameter search which has been tested so far if it has been initialised some distance from the actual minimum. Given that the cost function has been shown to be at a minimum when the data is correctly realigned (section 6.2.2), a necessary step in making the scheme practical for use on real data, which may contain inter-slice motion over several axes, will be to incorporate a more advanced optimisation framework. This optimisation method should be able to search the cost function space in a quasi-exhaustive manner so that the best minimum is always located.

A relatively simple improvement would seem to be to exploit the parallel structure of the permutation search in section 6.3.1 where, at present, 6 different permutations of the same parameters are optimised over the same data. By either retaining the direction-set optimisation already in place, or by switching to a gradient-descent scheme, the algorithm could be reimplemented to work on a parallel architecture. In this way, the multiple permutations could be evaluated concurrently which would mean that more time would be available for optimisation at a much finer level of detail than at present. The practical improvement in accuracy which it is hoped such a scheme will bring may be limited by the accuracy of the spatial interpolation and therefore by the accuracy of the cost function evaluation. Potential alternatives to the spatial reconstruction used in the existing implementation are discussed in section 7.1.1.

Yet another alternative, which may yield a more robust and accurate optimisation framework for the problem of spatio-temporal realignment, is optimisation by *simulated annealing*. This stochastic strategy has already been used successfully in image restoration [49], where the annealing process served to calculate the maximum *a posteriori* (MAP) estimates, given the original degraded images.

Simulated annealing uses a Metropolis procedure, originally adopted from statistical mechanics, in which controlled uphill steps can be incorporated into the optimisation, allowing the scheme to avoid local minima traps [88]. A further advantage of the scheme is that in the early stages of the optimisation, gross features of the eventual state of the system (in this case the orientation of the images) will dominate. As the optimisation progresses, however, fine details will begin to develop.

Typically, the annealing process is very computationally intensive, a situation which has motivated considerable research into adapting the method so that it can

be used in a realistic run-time. More recently, it has been shown that by making certain modifications to the annealing schedule, specifically by assuming a Cauchy distribution for the probability density of the state-space of the parameters [69], the method can find the minimum of a multi-dimensional optimisation problem several orders of magnitude more rapidly than other sophisticated techniques such as genetic algorithms [70].

It is recommended that any further work on spatio-temporal realignment of FMRI data, particularly involving images where the motion is not known *a priori*, should seek to implement the proposed TIGER model using a generalised simulated annealing method [137]. It should also be noted, however, that the current implementation takes approximately 72 hours to process a 20 volume time-series, including the permutation search, on a 64-bit ES40 667 MHz Alpha platform. This will scale linearly upwards in time when processing more typical data-sets which contain several hundred time-points. While accuracy is a priority, the practical limitations of a scheme which takes this long to run in a clinical setting must be considered. It may be possible to set aside to permutation search for each combination of rotation parameters. Despite the initial motivation for including this stage, all the test data produced nearly identical results, irrespective of the permutation of rotational search parameters. Abandoning this step will give a six-fold increase in speed, with several areas of the research code still suitable for optimisation.

7.1.3 Activation Regression

The problem of activation-led intensity changes driving the intensity-based registration has been brought up on a number of occasions throughout this thesis (see, for example, sections 3.3, 4.2). Recently, there has been considerable interest

in the design of cost functions which are robust to these experimentally-related changes within the context of rigid-body volumetric motion correction which is described below.

Robust similarity measures have been developed for FMRI data. These modify the least squares cost function with a Geman-McClure M-estimator [41] which de-weights the influence of intensity dissimilarities due to cerebral activation around areas of low gradient in the image but which enables areas of high gradient (for example around the edge of the brain) to play a far more influential role in the registration of the images. It was found that this method significantly improved the robustness of the cost function against activation-related confounding signal. It would seem to be straightforward to incorporate such a modification into the existing 4-D cost function used by the TIGER scheme.

It has been suggested [43] that by performing a two-stage motion correction, a preliminary activation detection can be carried out and the resulting voxels which have been determined to be associated with the experimental response can be excluded from a subsequent realignment step. If the nature of the experimental paradigm is known, it may be possible to regress out the expected form of the activation response from the data, leaving only baseline EPI signal. The registration is then far more likely to be robust to variations in image intensity which have not been caused either directly, or indirectly, by simple subject motion. In practice, this may be accomplished within the cost function by regressing out the activation signal from each time-course before carrying out variance calculations.

An integrated approach to rigid-body motion correction and activation detection has been developed [112], where the least squares operator between a floating image, represented by the matrix $\boldsymbol{F}$, and a reference image $\boldsymbol{G}$ is written as a first-order linear approximation:

$$\boldsymbol{F} = \boldsymbol{G} + \boldsymbol{A}\boldsymbol{X} + \boldsymbol{Y}\boldsymbol{B} \tag{7.1}$$

where $\boldsymbol{A}$ is a matrix where each row contains the gradient vector (with respect to each of the 6 rigid-body motion parameters) for a single voxel in $\boldsymbol{G}$, $\boldsymbol{X}$ contains a set of motion parameters corresponding to a different volume in the time-series on each row, $\boldsymbol{Y}$ contains the experimental regression parameters for each voxel/regressor combination[1] and $\boldsymbol{B}$ contains the regressors themselves. The iterative solution of

$$\min_{(\boldsymbol{X},\boldsymbol{Y})} ||\boldsymbol{A}\boldsymbol{X} + \boldsymbol{B}\boldsymbol{Y} - \boldsymbol{F} + \boldsymbol{G}|| \tag{7.2}$$

is then sought, where $||\cdot||$ refers to the Frobenius norm. Results using this scheme on simulated EPI data-sets showed that the integrated approach was better able to remove stimulus-correlated motion errors (see section 7.2.1) than conventional least squares motion correction, although some residual correlations remained in the data.

7.1.4 Intra-volume Motion

In the work to date it has been assumed that the subject motion during each stacked-slice volume acquisition can be modelled as a smoothly varying function [89] on the transformation parameters used to represent the overall motion. While convenient in terms of the complexity of the model which has been presented, it is physically plausible that the subject motion which takes place during an acquisition spaced over several seconds may in fact be predominantly discon-

[1] Where a regressor is some non-constant term which, when multiplied by an appropriate scalar, will produce a time-series that is close (in the least squares sense) to the actual data for a given voxel.

tinuous and have non-zero acceleration.

This may be exemplified by the case where motion occurs only towards either end of the acquisition so that for most of the time the patient is still ('brief twitch'); and equally in the case where there is continual motion throughout the scan but where the direction and magnitude of the motion may switch several times within a single volume. In each case, the existing model based on the screw decomposition of the initial affine estimate will not be able to detect such motion reliably and the subsequent incorrect estimate may be far from the true solution.

A solution to the current approach is to extend the initialisation stage to propose a number of motion 'velocity profiles' (figure 7.1), including the existing continuous model, but also including models which comprise the composition of two or more affine motions in succession. Any of these sub-motions may be zero to allow fitting the static case described in the 'brief twitch' example above. An optimisation or search can then be applied which presents each hypothesis and selects the model which leads to the lowest initial cost over 4-D.

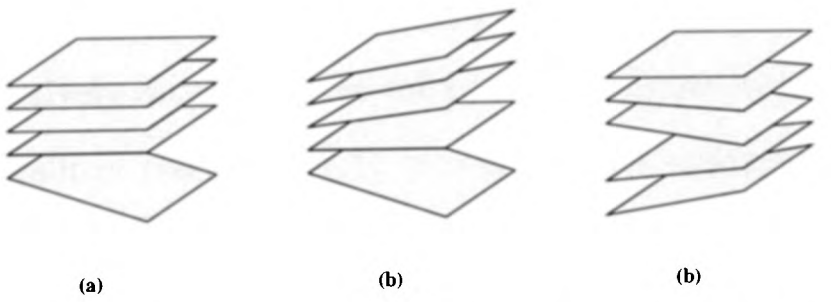


Figure 7.1: A wide variety of slice distributions is possible within each volume which may not be well described by the current assumption of smooth movement during acquisition. These alternative distributions include *(a)* sudden local movements; *(b)* increasing movement along a specified direction; and *(c)* a change of direction during the scan.

It must be borne in mind that such a scheme necessitates an increase of several orders of complexity to the problem, not least because it is necessary to evaluate a 4-D cost for each proposed velocity profile. In the case of profiles which contain

a step function, it will also be necessary to optimise within each volume to find the position of the step which provides the best fit to the overall cost function.

7.1.5 Multi-resolution Initialisation

Extending the multi-resolution scheme adopted in MCLFIRT, where a solution is sought over progressively finer levels of data resolution, a natural extension to the 4-D initialisation stage would be to search for the affine matrix which forms the basis of the initial fanned slice distribution using a coarse-to-fine strategy. This is particularly relevant to the spatio-temporal scheme, since the somewhat arbitrary decomposition of the estimated gross inter-volume motion into stepped slice planes is potentially highly inaccurate, given the limited detail in the original data. A further motivation, noted in section 6.2.1, is the fact that the initialisation stage contributes significantly to the quality of the overall correction as it implicitly performs much of the coarse-resolution search over the parameter space.

Rather than adopting a multi-scale approach just in the spatial domain, for example by recursively estimating the initial motion parameters through a range of isotropic re-scalings (section 3.1.4), a more robust strategy might be to allow the slice-scale to increase. In this way, the initial affine match for each pair of volumes $[V_n, V_{n+1}]$ would be allowed to decompose into a gradually increasing number of stepped slice matches between the volumes, each slice covering a greater number of voxels than is available at the actual slice resolution.

At present the initialisation scheme effectively moves from one slice (the entire volume) to the actual slice resolution without any intermediate steps. By introducing interim levels, the increased data within each larger slice enables the volumetric transformation to be refined into decomposed slice transformations.

This suggests a block matching search on the volume pair as depicted in figure 7.2.

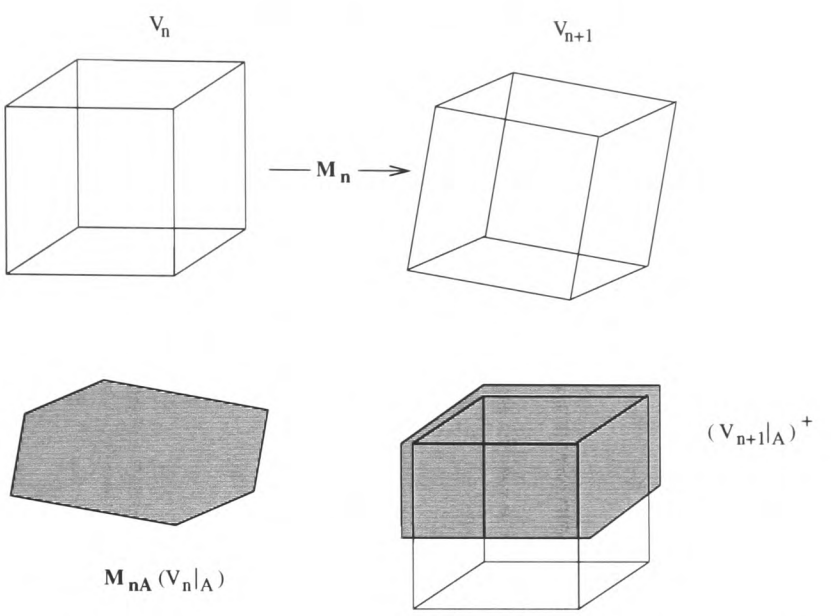


Figure 7.2: An initial volumetric (rigid-body) match $\boldsymbol{M_n}$ between two volumes, V_n and V_{n+1}, is decomposed into separate transformations, $\boldsymbol{M_{nA}}$ and $\boldsymbol{M_{nB}}$ which relate to 'sub-slices' A and B within V_n. These local-rigid transformations are refined using a block-matching scheme whereby a match is sought between each sub-slice and the dilation of the area in V_{n+1} (denoted $(...)^+$) described by its transformation under the appropriate sub-translation.

The refinement of the slice positioning can be achieved by partitioning the volumetric transformation $\boldsymbol{M_n}$ into separate transformations $\boldsymbol{M_{nA}}$ and $\boldsymbol{M_{nB}}$ using the screw coordinate decomposition of equation 5.1. The volumes can then be sub-divided in accordance with this decomposition so that a more local match can be sought. The recursive match could be achieved by dilating the target area[2]in the second volume $V_{n+1}|_z$ for each sub-volume $z = A, B$, and so on, and then matching the transformed slice from the target volume $\boldsymbol{M_{nz}}(V_n|_z)$ into this space. The potential for more accurate refinement of parameters using this approach is the main motivating factor but a potential confound may be that this increased flexibility renders the uniform fan-out invalid — it is no longer possible to base the position of a slice within a volume exclusively on the pose of

[2]Effectively increasing the area in V_{n+1} over which a match is computed with V_n.

its neighbours — and so a more advanced optimisation strategy which does not assume this smooth motion trend will be required.

7.1.6 Cost Function Design

A cost function, such as the one proposed in section 5.2.2, places equal weight on each voxel's ability to influence the overall cost: that is, the intensities are summed over all time-courses and then the total 4-D cost is compared with its previous value in the same way after each proposed perturbation to the slice distribution has been incorporated. A number of more subtle alternatives are available which may either affect the number of voxels over which the cost is computed or the types of perturbation which can be proposed by the scheme. For example, due to the largely homogeneous nature of many voxels at non-boundary regions within the brain, vertical translations typically lead to much smaller changes in an intensity-based cost function than those which result from rotations.

Knowledge of the likely distribution of features within the image (or at least an expected spatial intensity distribution if the low SNR of FMRI data is to be taken into account) might suggest an initial cost function computed using only boundary information, possibly using a parametrized spline model to describe the edges of the images [125]. Another common problem when estimating motion parameters is that rotations about one of the Cartesian axes are often hard to distinguish from translations along the other two orthogonal axes [30]. In these cases the cost function may include weighting factors, based on information of known ranges of physical movement in each of the parametrized directions, which attempt to drive the optimisation away from such potentially intractable scenarios. With these and other observations in mind it may be possible to design a cost function

optimisation in which larger perturbation steps are suppressed in directions which are expected to yield small cost function changes. Such an approach is inexorably linked with the design of the optimisation framework.

Of parallel consideration is the cost metric itself. In the interests of simplicity, work to date has centred on a Least Squares formulation, which has been shown to be both less accurate and more prone to false activations than correlation-based measures in a 3-D setting (section 3.2.1). The advantage of the current spatio-temporal cost function is that it can compactly describe the characteristics of the time-series as a variance measure. A development of other existing cost functions, most notably normalised correlation which gave robust and accurate results in the rigid-body framework, could be extended in a similar way since the correlation measure is also variance-based (see Appendix III). Given the complex evaluations already required by the current cost function, however, it is likely that a more sophisticated metric may well lead, in turn, to a much more difficult optimisation problem.

7.1.7 End-volume Optimisation

In the framework described in section 5.2.1, the slice distribution within a volume V_n depends on the pose of the previous volume V_{n-1} and subsequent perturbation and optimisation in 4-D (figure 5.17) assumes this flexibility for all volumes. This is not the case for the initial volume in the series whose first slice is always arbitrarily fixed but can be considered as a reference frame for the 4-D data. Ideally, the choice of initial slice pose in the first volume should not bias the overall solution. The fan distribution of slices (and the flexibility afforded to the end slice of a given volume and the first slice of the next volume) should, in turn, allow an accurate solution to be reached for each volume in the time-series

regardless of any preceding realignments.

Conversely, more flexibility but potentially less robustness is available to the last volume in the series, where the distribution of slices depends on the orientation of the last slice of the previous volume but in this case is not attenuated by the (non-existent) subsequent volume. This may lead to pathological situations in which the slice distribution across all but the final volume may be arbitrary, provided that the overall optimisation forces the slice distribution over the final volume to compensate for the increased cost. The sequential progression through the time-series, where all but the final volume are subject to the end slice constraint, should prevent such an error developing but the possibility of an inaccurate (but 'correct' in the sense of the local cost function minimisation) remains.

It may be that, in some cases, given the concerns above, the initialisation stage would provide more robust estimates if the order of the volumes were randomised to allow the end volumes the opportunity of local support available to other volumes in the time-series. As this will violate the assumption of small inter-volume movements with respect to adjacent volumes, more detailed study of the issues in this section would be warranted before any changes to the current implementation were to be implemented.

7.1.8 Extended Parameter Set

Thanks to the compact parameterisation of the existing scheme, there is scope to expand the number of degrees of freedom afforded to the optimisation at each time-point. There is likely to be a heuristic trade-off between increased flexibility and the robustness of the optimisation, given the caveats regarding fully independent slice movement described in section 2.3.5. With this in mind,

however, a logical development would be to allow both end slices, rather than simply the final slice of the test volume, to move (figure 5.17), and to compute the relative redistribution of slices between these boundaries. It should be borne in mind that there may be a number of redundant combinations of slice movements where more than one end slice is allowed to be perturbed, so any optimisation linked with this proposed method should seek to avoid such parameter choices.

Clearly, parameterisations other than a rigid-body 6 parameter set are possible but any change of this kind should be strongly suggested by acquisition modelling, rather than mathematical convenience, to ensure that the limited information inherent in the images is able to support the registration consistently.

7.1.9 Search Algorithm

Given the increased complexity of this scheme over standard rigid-body schemes, it may prove worthwhile to limit the time spent realigning individual volumes, particularly those which contain little relative motion artefact, by specifically seeking out those volumes which exhibit a greater level of motion corruption.

This might be achieved by pre-computing the initial variance cost of each volume with respect to the remainder of the time-series and then commencing the actual search by realigning the 'worst' volume given the criteria described above. It is likely, however, that it will be necessary to allow more lengthy optimisation of these outlier volumes which, by definition, will not correspond to the closely-matched data presupposed by the current initialisation step. Indeed, if it is likely that a volume being corrected does exhibit excessive inter-slice motion (characterised by a poor match to the variance of the remainder of the time-series data), it may be appropriate to increase the flexibility of the model used to realign this volume. Such an adaptation could incorporate some of the ideas presented

in sections 7.1.5, 7.1.4 and 7.1.8.

If it is possible to perform robust activation regression (section 7.1.3) it may also be possible to extend the parameter search so that the optimisation progresses at increasingly more refined levels of detail until a stipulated minimum cost value is reached, which may now be defined in the absence of any activation-related intensity outliers. This might ensure that a global minimum is more likely to be reached.

7.1.10 Spline Reconstruction

Clearly the choice of a Hermite spline basis (section 5.2.5) was but one of many possibilities, though it was motivated by its compact representation and the requirement that it should interpolate through its control points, that is, adjacent samples along the same time-course. Cubic B-spline basis functions have previously been used for temporal reconstruction in FMRI data [29] which suggests they are worthy of investigation in the context of spatio-temporal FMRI data reconstruction.

7.1.11 Temporal Interpolation

The main implementation issue which was described in section 5.2.5 is to expand the data storage used by the spatial interpolation within the optimisation scheme so that if a voxel is determined to have been acquired more than once, all the contributing voxel locations and their associated slice-timings are stored so that the final temporal correction can shift each contributing voxel before combining them. The other issue, as with many other more general interpolation concerns throughout the scheme, is whether or not the current trilinear and sinc

interpolation is the most suitable choice for the temporal re-shift. More accurate alternatives may be found by using spline kernels which can be designed to adjust for a range of irregular samples.

7.1.12 Future Developments

It has been observed that, as the resolution of FMRI images improves in the future, spatio-temporal corrections will become increasingly important [56], even for very small movements. At the same time, better corrections will be possible using the proposed TIGER model, in particular allowing more accurate interpolation and therefore more accurate cost function evaluations.

Alongside the expected maturing of FMRI technology, which is prerequisite to such improvements being made, developments in a number of other areas may be able to drive more accurate implementations of the TIGER model. One area in which this is most obvious is the issue of initialisation, introduced in figure 5.16 and developed in section 6.3. It is possible that it might be feasible to develop techniques which can robustly and consistently extract motion parameters from an ICA analysis, similar to the approach shown in figure 1.11. These estimates could be used to initialise the spatio-temporal optimisation.

An alternative means of obtaining these estimates might be through the use of external measuring devices. It is now possible to obtain accelerometers which can estimate rotation and translation at the level of millimetres. Such devices could be attached to the patient while he or she is in the scanner (possibly incorporated into the ear defenders which they are required to wear, and thus avoiding any additional invasive procedures) and be used to provide additional inputs to the optimisation initialisation. Three immediate areas of difficulty with such an approach spring to mind, however: the scheme clearly cannot be applied

to data-sets which have already been obtained, limiting the usefulness of the spatio-temporal correction to new data alone; the cost of such a device may prove prohibitive[3]; and the issues involved with using a gyroscopic device inside a strong magnetic field may well be non-trivial.

Once possible future directions have been outlined in the next section, a number of further areas of investigation present themselves. Some of these relate to potential improvements to the model and its implementation developed in the previous chapter while other issues, which up until now have not been afforded much attention by the FMRI community in general, arise as a result of the increased potential accuracy available with this new model.

7.2 Open Issues

Almost paradoxically, one side-effect of providing a correction for a specific artefact such as motion within an intrinsically noisy modality such as FMRI is that more subtle effects begin to dominate the data. In this section, two of the more widely-recognised of these effects are considered: stimulus-correlated motion and magnetic spin history. These can both be considered as second order motion-related confounds and it seems appropriate to suggest ways in which the proposed method may be adapted to model and correct for the distortions to which they invariably lead.

7.2.1 Stimulus-correlated Motion

Motion which occurs during the acquisition of FMRI data can be separated broadly into two groups: stimulus-correlated and uncorrelated. Stimulus-correlated

[3] At the time of writing, a reasonably accurate small gyro costs a few hundred pounds, and an accelerometers about one hundred pounds.

motion is most common in studies where the experiment being carried out necessitates a physical reaction to a stimulus, for example a motor experiment where the subject has been instructed to move his or her fingers in response to a visual cue.

In the context of the General Linear Model described by equation 1.12, motion artefacts (both stimulus-correlated and uncorrelated) have adverse effects on the estimation of the response to a stimulus. In the case of uncorrelated motion, if the motion is not modelled as a specific effect in the GLM analysis, the variance assumed to be attributable to noise will be overestimated. While the estimates of the activation response parameters ($\boldsymbol{B}$ in equation 1.12) remain unchanged (compared with the non-confounded case),[4] the inflated noise standard error will lead to a lower z-statistic. In the case where there is a correlation between motion parameters and the experimental stimulus (which is nearly always at least partially the case [57]), the $\boldsymbol{B}$ values will either be overestimated for a positive correlation, or underestimated if there is a negative correlation, between the stimulus response and the motion. Either situation will lead to inaccurate final z-statistics [17] and be potentially dangerous in a clinical setting.

The alternative approach to this analysis, whereby the movement effects are modelled within the analysis, typically by including the estimated motion parameters as regressors in the GLM analysis, can lead to more accurate results if the motion and the experimental response are completely orthogonal to one another. In the more likely case of (at least partial) correlation, however, a decision will have to be made as to whether to attribute the explained signal exhibited both by motion and the experimental regressor to one or the other source.

ICA is able to detect the presence of stimulus-correlated motion by separat-

[4] Assuming that autocorrelation is ignored.

ing the activation and the motion [6] as shown in figures 7.3 and 7.4 respectively. In this example, an ICA was carried out on data from a boxcar design visual experiment and identified two independent components whose characteristic frequency matched that of the stimulus. One of these can be attributed to the visual stimulus while the other is labelled as stimulus-correlated motion given that the corresponding spatial map exhibits the same characteristics seen in other motion artefacts, such as those in figures 1.11 where there is a strong intensity shift along the edge of the brain.

Attempts have been made to automate the detection of stimulus-correlated motion with ICA and remove the components corresponding to this effect [90] but the restrictions to this method are such that it can only be applied with any accuracy to a very limited set of data.

7.2.2 Spin History

If the results of the ICA on motion corrected real data are examined, it is possible to identify what appear to be other known confounds in the data. One common confound, which is rarely compensated for in FMRI processing, is the spin history effect. Because it is assumed that the system of spins within each object-referenced slice will be excited by separate RF pulses at regular intervals (section 1.2.1), the steady-state magnetisation of the tissue is disrupted when an object moves in the scanner between the acquisition of individual slices and it will require a number of acquisitions before the longitudinal magnetisation returns to steady state. It follows that, if the spin system has not returned to equilibrium before the next RF excitation pulse is applied, the magnetic state of the system will depend on the history of past magnetic states [110].

In figure 7.5 the temporal derivative of the estimated z-translation found

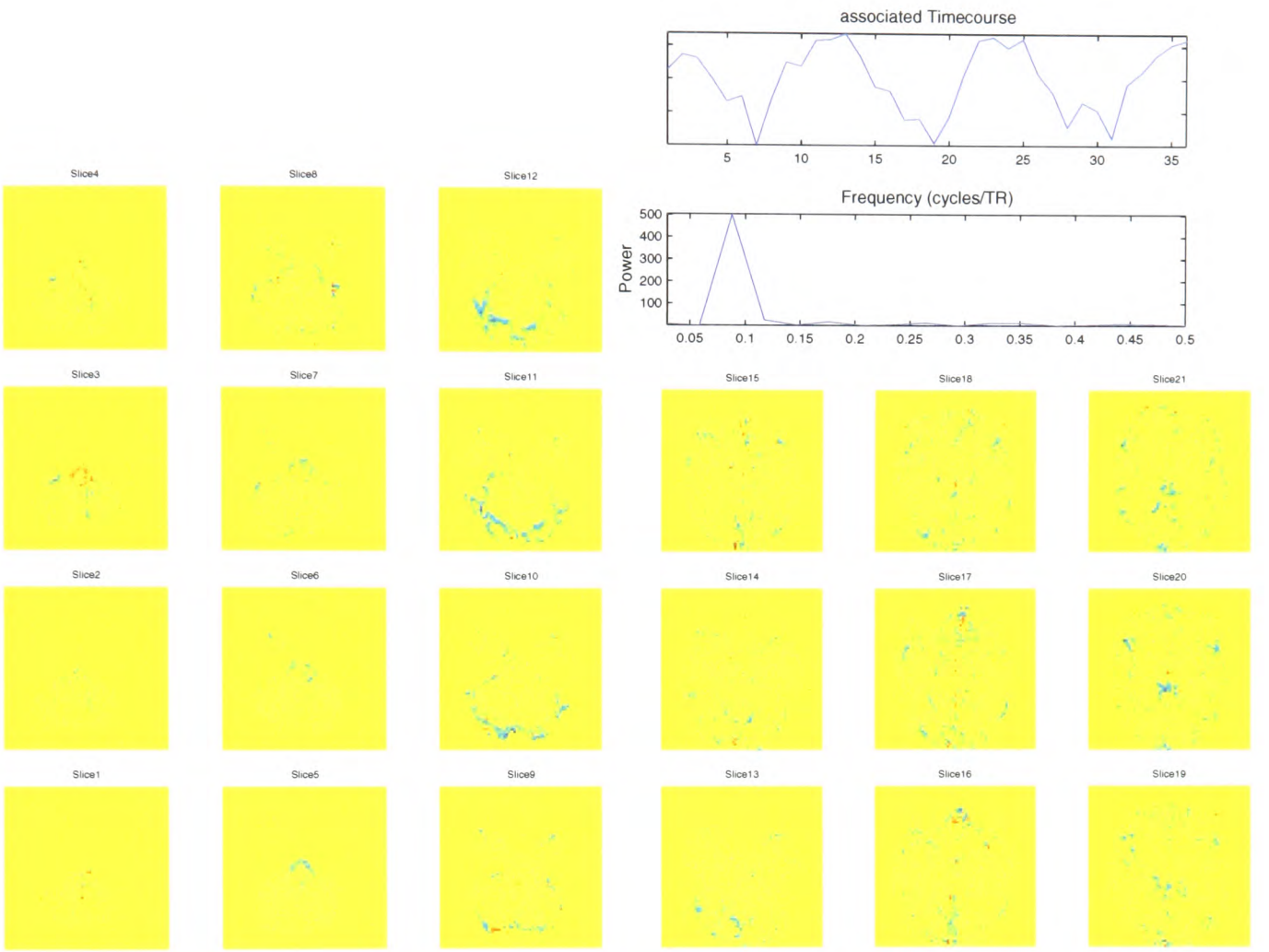


Figure 7.3: ICA time-courses and spatial maps corresponding to the experimental visual paradigm for a study containing stimulus-correlated motion. This component is associated with the experimental response (that is, the signal of interest) and has been labelled as such due to the strongly periodic temporal profile in the top right-hand plot (confirmed by the FT shown below) and some peak activation (marked as red) in the spatial maps. Note, however, that the spatial localisation is much less clear than in experiments that are not thought to contain stimulus-correlated motion, for example, figure 3.14*(bottom)*.

using motion correction has been plotted against an extracted IC time-course. The strong and consistent correspondence between peaks in both plots supports the conclusion that this IC is associated with spin history effects which themselves are functions of the changes in slice position and, hence, timing.

Spin history effects are due to incomplete M_z recovery as a result of through-plane motion within a magnetic field, and as such they are prime candidates for motivating a correction scheme which adaptively updates the scanner position to

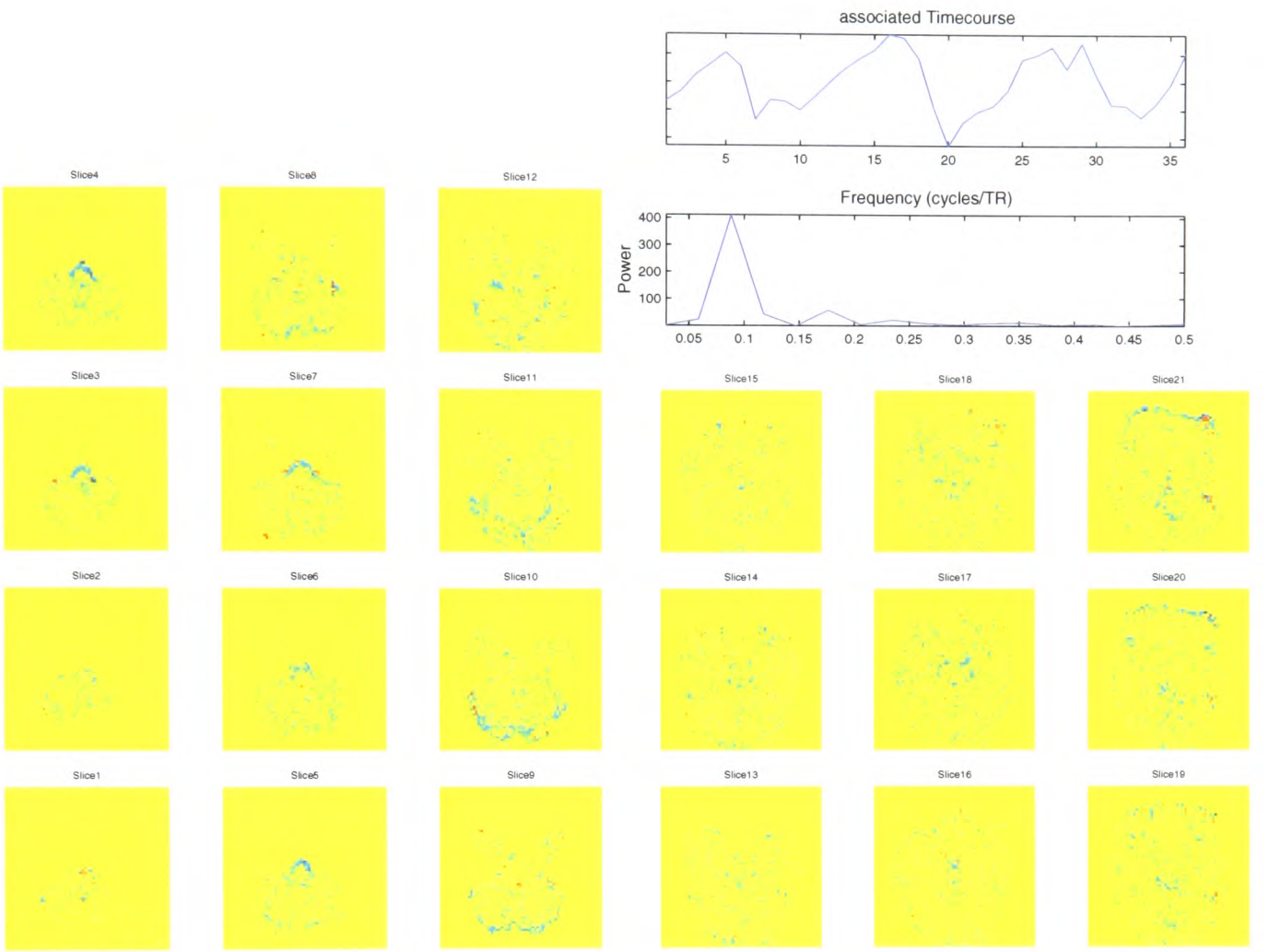


Figure 7.4: ICA time-courses and spatial maps corresponding to the stimulus-correlated motion artefact for the study in figure 7.3. Comparing the temporal characteristic with the previous figure, it is hard to discriminate clearly between the experimental response and the motion artefact shown in this figure. Labelling is performed by examining the spatial components which, in this figure, show edge intensity effects such as those shown in figure 7.5.

take into account any movement by the subject and to ensure that, over the scan period, the slice selection remains fixed with respect to the subject.

Friston *et al.* [46] proposed a method of post-alignment intensity correction based on the observation that the number of excited spins (and therefore the signal) will be a function of position in current and previous scans. It was noted that the spin history effect will manifest itself if the recovery of magnetisation in the z direction is incomplete by the time the next slice-selective pulse arrives (that is, if $TR < T_1$ which is likely in many FMRI studies).

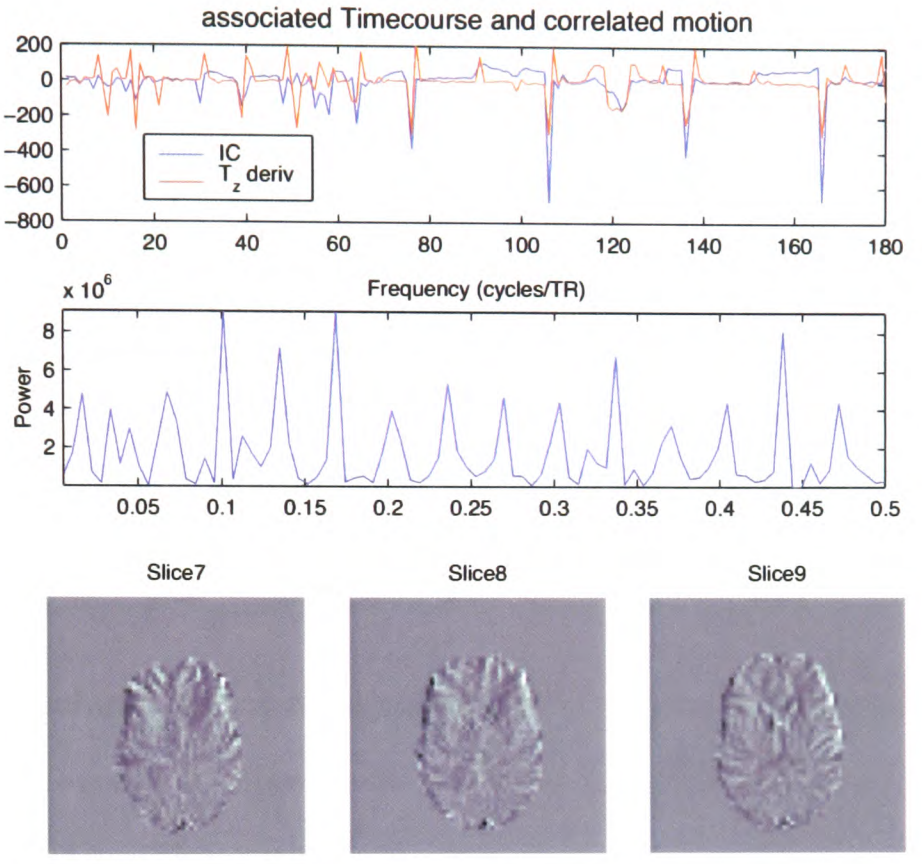


Figure 7.5: IC time-course (with scaled derivative of z-translational motion overlaid) in top plot and selected spatial maps in bottom plot showing spin history effect. Spin history is a second-order motion effect so it is expected that it will be temporally correlated with the derivative of the patient motion parameters. Spatially, the effect is similar to rigid-body motion with ridges of markedly increased or decreased intensity along the edge of the brain. This can be seen as the black and white artefacts in the lower figure.

Assuming a rigid-body transformation, a Least Squares estimate is formed on the 6 motion parameters by taking the first order Taylor expansion (for small transformations) with the object and reference images written as column vectors and represented as scalar functions of position in space. The next step is to partition the measured signal from each voxel into orthogonal components where one component represents the artefact-free signal of interest and the other component is the assumed functional component of the measured signal due to motion (including motion history) effects.

By modelling the proportional reduction in $M_z(t)$ after excitation, where the

reduction is a measure of the relative degree of excitation elicited by the RF pulse, the second component can be modelled as the sum of second order polynomials over the components of the displacement, where the choice of second order relates to the possibility that positional dependence of relative excitation will be curvilinear, that is, movement in z out of a slice in either direction will reduce the excitation. This gives the correction required for each voxel time-course $\mathbf{X}$ as:

$$\mathbf{X}^* = \mathbf{X} - \mathbf{Q} \cdot (\mathbf{Q}^\mathbf{T} \cdot \mathbf{Q})^{-1} \cdot \mathbf{Q}^\mathbf{T} \cdot \mathbf{X} \tag{7.3}$$

where $\mathbf{X}^*$ is a column vector of *adjusted* FMRI values for the voxel in question completely orthogonal to movement artefacts modelled in terms of the movement estimates $\mathbf{Q}$. The authors observe that the smoothness of the images should be greater than the relative displacements between them (enforced by pre-convolution with a Gaussian filter). They also comment that the method is unsuitable for PET studies where the number of scans is less than the number of columns of $\mathbf{Q}$.

What the SPM approach [46] aims to achieve is to remove the effects of motion that are caused by movement during the acquisition at a time-point distinct from the time-point of the current volume under consideration. Equation 7.3 represents the signal of interest $\mathbf{X}^*$ when all the motion-related confounds have been removed from the raw data $\mathbf{X}$. The authors were able to demonstrate that by applying this signal correction it was possible to demonstrate significant activation in a study where the activation had eluded detection, even with conventional motion correction.

Later research [123], however, looked at motion in the through-plane direction

which demonstrated that the correlation due to saturation is proportional to Δ^2, where Δ is the range of motion between successive slices, and the effects of structure increase linearly with z. The authors concluded that the source of intensity fluctuation due to the variations in saturation is extremely small and it may be beneficial to ignore this effect. They note that until better registrations are available, the variance can be reduced by removing correlations with z and z^2 and perhaps x and y.

These findings have meant that more recent versions of the SPM package have made the correction of equation 7.3 obsolete [93]. It was also observed that because the spin history effect is correlated to z-translation, removal of any motion correlated components could lead to the introduction of additional artefacts in the data. This does not negate the validity of spin history correction but highlights the need for a more accurate model which can be developed to allow appropriate corrections in the data.

More recently, new quantitative assessments have been carried out to determine how important the spin history effect is in the correction of time-series [110]. Returning to the observation that if all the regions of the object are in magnetic equilibrium prior to each volume acquisition, there will be no spin history effect, the authors of this research investigated the effect of head motion between successive volume acquisitions on the steady-state magnetisation of the scanned object. It was demonstrated that a back and forth motion of the head during a single volume acquisition destroyed the steady state signal for several subsequent volumes, leading to a signal change of 3% to 7%. Thus the issue of spin history correction is again a focus of attention in the FMRI community.

7.2.3 Field Corrections

One final area of artefact correction which is strongly linked with the issue of motion is that of susceptibility distortions. These were introduced in section 1.4.3, where it was shown that the presence of a distinct tissue interface (specifically the air-filled sinuses in the frontal regions of the skull) could interact with the applied magnetic field to yield non-linear distortions. By placing an object (for example, the head) into the magnetic field of the scanner, gradients are introduced which disturb the local flux field. This leads to problems with the reconstruction applied to the data which relies on gradient information (section 1.2.1). The gradient will now be equal to the sum of the gradient of the coil *and* the gradient introduced by the subject and distortions will manifest in the phase encode direction.

These distortions will vary with the position of the head in the scanner and therefore with motion. As such, the spatio-temporal registration scheme is ideally placed to incorporate recently developed approaches to correcting these distortions. The approach taken in [1] is to model the change in magnetic (B_0) field as a 1st order Taylor expansion on the motion parameters and then solve this using a least-squares approach. Results have shown that this approach is able to give qualitatively good agreement with externally-measured parameters.

An alternative approach, developed in [76], is to compute explicitly the change in B_0 as a function of position [27].

By including one of these methods in the intensity estimation and interpolation scheme used by the realignment, it may be possible to incorporate these corrections into the model-based framework proposed in Chapter 5.

Appendix I

GLM Estimation

Referring back to equations 1.12 and 1.13, there exists [127] a square, non-singular matrix $\mathbf{K}$ such that $\mathbf{V} = \mathbf{K}\mathbf{K}^{\mathbf{T}}$, and that $\mathbf{e} = \mathbf{K}\epsilon$ where ϵ are $MVN(0, \sigma^2\mathbf{I})$.

Now consider a GLM which incorporates a general linear temporal filtering of the data, where $\mathbf{S}$ is the square $T \times T$ matrix that performs the temporal filtering via matrix multiplication. $\mathbf{S}$ is a Toeplitz matrix produced from the impulse response; this is directly equivalent to convolving with the impulse response using zero padding. The design matrix is also temporally filtered using $\mathbf{S}$ to reflect the known change in the observed data, giving

$$\mathbf{SY} = \mathbf{SXB} + \eta \tag{I.1}$$

where η is $MVN(0, \sigma^2\mathbf{SVS}^{\mathbf{T}})$. An ordinary least squares (OLS) estimate of $\mathbf{B}$ is used, given by:

$$\widehat{\mathbf{B}} = (\mathbf{SX})^{+}\mathbf{SY} \tag{I.2}$$

where $(\mathbf{SX})^{+}$ is the pseudo-inverse of $(\mathbf{SX})$ given by

$$(\mathbf{SX})^{+} = ((SX)^{T}SX)^{-1}(SX)^{T}. \tag{I.3}$$

The variance of a contrast $\mathbf{c}$, of these parameter estimates, $\widehat{\mathbf{B}}$, is given by:

$$\begin{aligned} Var\{\mathbf{c^T}\widehat{\mathbf{B}}\} &= k_{eff}\sigma^2 \\ k_{eff} &= \mathbf{c^T}(\mathbf{SX})^{+}\mathbf{SVS^T}((\mathbf{SX})^{+})^{\mathbf{T}}\mathbf{c} \end{aligned} \tag{I.4}$$

To use equation I.4 an estimate of σ^2 [150, 127] is required:

$$\widehat{\sigma}^2 = \eta^{\mathbf{T}}\eta/\mathbf{trace}(\mathbf{RSVS^T}) \tag{I.5}$$

where $\mathbf{R} = \mathbf{I} - \mathbf{SX}(\mathbf{SX})^{+}$, the residual forming matrix, which can be used to obtain the residuals of the model fit:

$$\mathbf{r} = \mathbf{RSY} \tag{I.6}$$

Appendix II

Independent Component Analysis

Classical ICA expresses the data from an FMRI experiment with n voxels measured at p different time points as a $p \times n$ matrix $\mathbf{X}$ for which a decomposition is sought such that

$$\mathbf{X} = \mathbf{AS} \tag{II.1}$$

where the matrix $\mathbf{S}$ is optimised to contain statistically independent maps in its rows, i.e. spatial areas in the brain, each with an internally consistent temporal dynamic which is characterised by a time-course contained in the associated column of the square *mixing matrix* $\mathbf{A}$. The sources are estimated by iteratively optimising an *unmixing matrix* $\mathbf{W} = \mathbf{A}^{-1}$ using the *infomax* algorithm [10] or similar so that $\mathbf{S} = \mathbf{WX}$ contains mutually independent rows.

While the ICA model described above is a simple linear model, it differs from the standard GLM used in neuroimaging in two aspects: the mixing is assumed to be square[1] and the model of equation II.1 does not include a noise model,

[1] that is, there are as many noise sources as there are time-points.

the latter condition precluding the assessment of statistical significance of the source estimates within the framework of null-hypothesis testing. This leads to the problem of over-fitting a noise-free generative model to noisy observations which in turn suggests the use of a suitable probabilistic model within the ICA framework that controls the balance between what is attributable to 'real effects' of interest and what is simply due to observational noise. These issues are addressed through the use of a probabilistic ICA (PICA) model [118, 8] for FMRI data that allows for a non-square mixing process and assumes that the data are confounded by additive Gaussian noise.

Appendix III

Cost Functions

The following cost functions are commonly used in medical image registration as a measure of goodness of fit between two images. The two images are usually referred to respectively as the 'test image' and the 'reference image' where an optimal alignment of the test image is sought with respect to the reference image.

Least Squares

This is the simplest cost function which measures the sum of squared differences between voxel pairs in a reference image I_{ref} and I_{test}:

That is, let X represent the set of all voxel intensities in I_{ref} and Y represent the set of all voxel intensities in $T(I_{test})$. The normalised correlation is then defined as:

$$LS = \sum (Y - X)^2 \tag{III.1}$$

Normalised Correlation

Given a reference image I_{ref} and a transformed test image $T(I_{test})$, the normalised correlation is the correlation over the set of all voxel intensities in the two images,

written X and Y respectively.

$$NC = \frac{Cov(XY)}{\sqrt{(Var(X) * Var(Y))}} \tag{III.2}$$

where $Var(...)$ denotes the variance.

Mutual Information (MI)

Intuitively, if two images are correctly matched, knowing one image gives information about the other. Therefore their *mutual information* is high. Conversely, two independent signals or images will have a very low mutual information. If we let X be a random variable (RV), $P(X)$ be the probability distribution of X and $p(x)$ be the probability density of X then the entropy of X, $H(X)$, is defined as:

$$H(X) = -E_X\left[\log(p(X))\right] \tag{III.3}$$

Entropy is a measure of randomness. The more random a variable is, the more entropy it will have. The joint entropy is a statistic that summarises the degree of dependence of a RV X on an other RV Y. It is defined by:

$$H(X,Y) = -E_X\left[E_Y\left[\log(p(X,Y))\right]\right] \tag{III.4}$$

The conditional entropy is a statistic that summarises the randomness of Y given knowledge of X. It is defined by:

$$H(X|Y) = -E_X\left[E_Y\left[\log(p(X|Y))\right]\right] \tag{III.5}$$

Two random variables are considered to be independent if:

$$H(X,Y) = H(X) + H(Y) \tag{III.6}$$

The Mutual Information, MI, between two random variables X and Y is given by:

$$MI(X,Y) = H(Y) - H(Y|X) = H(X) + H(Y) - H(X,Y) \tag{III.7}$$

Thus it is a measure of the reduction of the entropy of Y given X [50].

Normalised Mutual Information [133] is simply given as:

$$NMI(X,Y) = \frac{H(X,Y)}{H(X) + H(Y)} \tag{III.8}$$

Correlation Ratio

The correlation ratio requires that the reference image be partitioned into a number of intensity isosets, that is, broken up into areas of similar intensity. The boundaries of these areas are placed over the (transformed) test image. Then the variance within each area is calculated and the cost is defined as a weighted sum of the variances divided by a normalisation term.

$$CR = \frac{\sum_i w_i Var(S_i)}{Var(S)} \tag{III.9}$$

where S_i is the ith isoset and $w_i = \frac{N_i}{N}$ are weighting terms, where N_i is the number of voxels in S_i and N is the number of voxels in S. Note that the normalisation is global.

Ratio of Image Uniformity

This measure was developed by Woods *et al.* [146] and also requires intensity partitioning of the reference image. It can be expressed as the weighted mean normalised standard deviation of the image intensities:

$$RIU = \sum_i \frac{N_i}{N} \frac{\sqrt{\mathrm{Var}(S_i)}}{\mu(S_i)} \tag{III.10}$$

where $\mu(\ldots)$ is the mean.

Bibliography

[1] J.L.R. Andersson, C Hutton, J Ashburner, R Turner, and K Friston. Modelling geometric distortions in EPI time series. *NeuroImage*, 2001.

[2] J. Ashburner, P. Neelin, D.L. Collins, A. Evans, and K. Friston. Incorporating prior knowledge into image registration. *NeuroImage*, 6:344–352, 1997.

[3] D. Atkinson, D.A. Porter, D.L.G. Hill, F. Calamante, and A. Connelly. Sampling and reconstruction effects due to motion in diffusion-weighted interleaved echo planar imaging. *Magnetic Resonance in Medicine*, 44(1):101–109, 1999.

[4] R. Bajcsy and S. Kovacic. Multi-resolution elastic matching. *Computer Vision, Graphics, and Image Processing*, 46(1):1–21, 1989.

[5] P.R. Bannister, C.F. Beckmann, and M. Jenkinson. Exploratory motion analysis in fMRI using ICA. In *Seventh Int. Conf. on Functional Mapping of the Human Brain*, 2001.

[6] P.R. Bannister, C.F. Beckmann, and M. Jenkinson. Motion artefact decorrelation in FMRI analysis using ICA. In *International Soc. of Magnetic Resonance in Medicine*, 2002.

[7] P.R. Bannister and M. Jenkinson. Robust affine motion correction in fMRI time series. In *Seventh Int. Conf. on Functional Mapping of the Human Brain*, 2001.

[8] C.F. Beckmann, J.A. Noble, and S.M. Smith. Investigating the intrinsic dimensionality of FMRI data for ICA. In *Seventh Int. Conf. on Functional Mapping of the Human Brain*, 2001.

[9] C.F. Beckmann and S.M. Smith. Probabilistic ICA for FMRI — noise and inference. In *Fourth Int. Symp. on Independent Component Analysis and Blind Signal Separation*, 2003.

[10] A.J. Bell and T.J. Sejnowski. An information-maximisation approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–1159, 1995.

[11] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, 3:259–302, 1986.

[12] P.J. Besl and N.D. McKay. A method for registration of 3-D shapes. *PAMI*, 14(2):239–256, February 1992.

[13] B.B. Biswal and J.S. Hyde. Contour-based registration technique to differentiate between task-activated and head motion-induced signal variations in fMRI. *Magnetic Resonance in Medicine*, 38:470–476, 1997.

[14] A. Blake and M. Isard. *Active Contours: The Application of Techniques from Graphics, Vision, Control Theory and Statistics to Visual Tracking of Shapes in Motion.* Springer-Verlag, 1998.

[15] M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. In *Proc. Visualization in Biomedical Computing*, volume 1131 of *Springer Lecture Notes in Computer Science*, pages 267–276, Hamburg, 1996.

[16] C. Broit. *Optimal Registration of Deformed Images.* PhD thesis, University of Pennsylvania, August 1981.

[17] E.T. Bullmore, M.J. Brammer, S. Rabe-Hesketh, V.A. Curtis, R.G. Morris, S.C.R. Williams, T. Sharma, and P.K. McGuire. Methods for diagnosis and treatment of stimulus-correlated motion in generic brain activation studies using fMRI. *Human Brain Mapping*, 7:38–48, 1999.

[18] T. Buzug and J. Weese. Improving DSA images with an automatic algorithm based on template matching and an entropy measure. *Computer Assisted Radiology*, 1124:145–150, 1996.

[19] P. Cachier, X. Pennec, and N. Ayache. Fast non rigid matching by gradient descent: Study and improvements of the demons algorithm. Technical report, INRIA, June 1999.

[20] P.T. Callaghan. *Principles of Nuclear Magnetic Resonance Microscopy.* Clarendon Press, Oxford, 1991.

[21] G.E. Christensen. *Deformable shape models for anatomy.* PhD thesis, Washington University, August 1994.

[22] G.E. Christensen, R.D. Rabbitt, and M.I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions of Image Processing*, 5(10):1435–1447, 1996.

[23] S. Clare. *Functional MRI - Methods and Applications.* PhD thesis, Nottingham University Department of Physics, 1997.

[24] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality registration based on information theory. In Y Bizais, C Barillot, and R Di Paolo, editors, *Proceedings of Information Processing in Medical Imaging*, pages 263–274. Kluwer Academic, 1995.

[25] D.L. Collins, P. Neelin, T.M. Peters, and A.C. Evans. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, 18(2):192–205, 1994.

[26] G.S. Cox and G. de Jager. Automatic registration of temporal pairs for digital subtraction angiography. *Medical Imaging: Image Processing*, 2167:188–199, 1994.

[27] R.W. Cox. Motion and Functional MRI: Informal notes for the Boston '96 workshop on functional MRI, 1996.

[28] R.W. Cox and A. Jesmanowicz. Real-time 3D image registration for functional MRI. *Magnetic Resonance in Medicine*, 42:1014–1018, December 1999.

[29] N. Crellin, T. Hastie, and I. Johnstone. Statistical models for image sequences. Technical report, Stanford University, 1998.

[30] K. Daniilidis and H.H. Nagel. The coupling of rotation and translation in motion estimation of planar surfaces. In *Conference on Computer Vision and Pattern Recognition*, pages 188–193, 1993.

[31] E. De Castro and C. Morandi. Registration of translated and rotated images using finite Fourier transforms. *IEEE Trans. PAMI*, 5:700–703, 1987.

[32] J. Declerck, G. Subsol, J.-P. Thirion, and N. Ayache. Automatic retrieval of anatomical structures in three-dimensional medical images. Technical Report 2485, INRIA, 1995.

[33] R. Duda and P. Hart. *Pattern Classification and Scene Analysis.* John Wiley and Sons, 1973.

[34] G.J. Ettinger, W.E.L. Grimson, M.E. Leventon, R. Kikinis, V. Gugino, W. Cote, M. Karapelou, L. Aglio, M. Shenton, G. Potts, and E. Alexander. Non-invasive functional brain mapping using registered transcranial magnetic stimulation. *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, 1996.

[35] A.C. Evans, W. Dai, D.L. Collins, P. Neelin, and S. Marrett. Warping of a computerized 3D atlas to match brain image volumes for quantitative neuroanatomical and functional analysis. *SPIE Medical Imaging*, 1445:236–247, 1991.

[36] J. Feldmar and N. Ayache. Rigid, affine and locally affine registration of free-form surfaces. *International Journal of Computer Vision*, 18:99–119, 1996.

[37] M. Firbank, C.E. Elwell, C.E. Cooper, and D.T. Delpy. Experimental and theoretical comparison of NIR spectroscopy methods for measuring cerebral haemoglobin changes. *J. Appl. Physiol.*, 85(1915–1921), 1998.

[38] D.B. Fitzgerald, G.R. Cosgrove, S. Ronner, H. Jiang, B.R. Buchbinder, J.W. Belliveau, B.R. Rosen, and R.R. Benson. Location of language in the cortex: A comparison between functional MR imaging and electrocortical stimulation. *American Journal of Neuroradiology*, 19:1529–1539, 1997.

[39] J.D. Foley, A. van Dam, S.K. Feiner, and J.F. Hughes. *Computer Graphics: Principles and Practice.* Addison Wesley, 2nd edition, 1996.

[40] W. T. Freeman and E. C. Pasztor. Markov networks for super-resolution. In *34th Annual Conference on Information Sciences and Systems*, 2000.

[41] L. Freire and M. Jenkinson. A gradient-informed robust motion correction method for FMRI. In *International Workshop on Biomedical Image Registration*, 2003.

[42] L. Freire and J.-F. Mangin. Motion correction algorithms of the brain mapping community create spurious functional activations. In M.F. Insana and R.M. Leahy, editors, *17th Int. Conf. on Information Processing in Medical Imaging*, pages 246–258. Springer, June 2001.

[43] L. Freire and J.-F. Mangin. Two-stage alignment of fMRI time series using the experiment profile to discard activation-related bias. In *Medical Image Computing and Computer-Assisted Intervention*, Lecture Notes in Computer Science, pages 663–670. Springer, 2002.

[44] K.J. Friston, J. Ashburner, C.D. Frith, J.-B. Poline, J.D. Heather, and R.S.J. Frackowiak. Spatial registration and normalization of images. *Human Brain Mapping*, 2:165–189, 1995.

[45] K.J. Friston, P. Jezzard, and R. Turner. Analysis of functional MRI time-series. *Human Brain Mapping*, 1:153–171, 1994.

[46] K.J. Friston, S.R. Williams, R. Howard, R.S.J. Frackowiak, and R. Turner. Movement-related effects in fMRI time-series. *Magnetic Resonance in Medicine*, 35:346–355, 1996.

[47] Z.W. Fu, Y. Wang, R.C. Grimm, P.J. Rossman, J.P. Felmlee, S.J. Riederer, and R.L. Ehman. Orbital navigator echoes for motion measurements in magnetic resonance imaging. *Magnetic Resonance in Medicine*, 34:746–753, 1995.

[48] T.S. Geisel and O. Acosta. Dammit, Smithers! This isn't rocket science, it's brain surgery! In *Proc. Samoan Conference on Simian Neuroscience*, pages 1938–1990, 2003.

[49] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.

[50] S. Gilles. Description and experimentation of image matching using mutual information. Technical report, Robotics Research Group, Department of Engineering Science, Oxford University, 1996.

[51] G.H. Glover. 3D z-shim method for reduction of susceptibility effects in BOLD fMRI. *Magnetic Resonance in Medicine*, 42:290–299, 1999.

[52] D.M. Grant and R.K. Harris, editors. *Encyclopedia of Nuclear Magnetic Resonance*, volume 1: Historical Perspectives. Wiley, 2002.

[53] S. Grootoonk, C. Hutton, J. Ashburner, A. Howseman, O. Josephs, G. Rees, K.J. Friston, and R. Turner. Characterization and correction of interpolation effects in the realignment of fMRI time series. *NeuroImage*, 2000.

[54] R. Guillemaud and J.M. Brady. Enhancement of MR images. In *Visualisation in Biomedical Computing (LNCS:1131)*, pages 107–116, 1996.

[55] A. Guimond, A. Roche, N. Ayache, and J. Meunier. Multimodal brain warping using the demons algorithm and adaptive intensity corrections. Technical Report 3796, INRIA, November 1999.

[56] J.V. Hajnal, D.L.G. Hill, and D.J. Hawkes, editors. *Medical Image Registration.* CRC Press, 2001.

[57] J.V. Hajnal, R. Myers, A. Oatridge, J.E. Schwieso, I.R. Young, and G.M. Bydder. Artefacts due to stimulus correlated motion in functional imaging of the brain. *Magnetic Resonance in Medicine*, 31:283–291, 1994.

[58] J.V. Hajnal, N. Saeed, A. Oatridge, E.J. Williams, I.R. Young, and G.M. Bydder. Detection of subtle brain changes using subvoxel registration and subtraction of serial MR images. *Journal of Computer Assisted Tomography*, 19(5):677–691, 1995.

[59] J.V. Hajnal, N. Saeed, E.J. Soar, A. Oatridge, I.R. Young, and G.M. Bydder. A registration and interpolation procedure for subvoxel matching of serially acquired MR images. *Journal of Computer Assisted Tomography*, 19(2):289–296, 1995.

[60] P.M. Hayton, M. Brady, S.M. Smith, and N. Moore. A non-rigid registration algorithm for dynamic breast MR images. *Artificial Intelligence*, 114(1):125–156, October 1999.

[61] P.F. Hemler, T.S. Sunamaweera, P.A. van den Elsen, S. Napel, and J.R. Adler. A versatile system for multimodality image fusion. *Journal of Image Guided Surgery*, 1:35–45, 1995.

[62] N. Herodotou, A.N. Venetsanopoulos, and L. Onural. Image interpolation using a simple Gibbs random field model. In *International Conference on Image Processing*, volume 1, pages 494–497, 1995.

[63] D.L.G. Hill. *Combination of 3D Medical Images from Multiple Modalities.* PhD thesis, University of London, 1993.

[64] D.L.G. Hill. From medical images and signals to clinical information: An interdisciplinary research collaboration proposal. `http://www.robots.ox.ac.uk/~irc`, 2001.

[65] D.L.G. Hill and D.J. Hawkes. Voxel similarity measures for automated image registration. In *Visualization in Biomedical Computing*, volume 2359, pages 205–216. Proc. SPIE, 1994.

[66] D.L.G. Hill, D.J. Hawkes, N. Harrison, and C.F. Ruff. A strategy for automated multimodality registration incorporating anatomical knowledge and imager characteristics. In H.H. Barrett and A.F. Gmitro, editors, *Information Processing in Medical Imaging*, pages 182–196. Springer-Verlag, 1993.

[67] C.K. Hoh, M. Dahlbom, G. Harris, Y. Choi, R.A. Hawkins, M.E. Phelps, and J. Maddahi. Automated iterative three-dimensional registration of positron emission tomography images. *Journal of Nuclear Medicine*, 34:2009–2018, 1993.

[68] B.K.P. Horn and B.G. Schunck. Determining optical flow. *Articial Intelligence*, 17:185–203, 1981.

[69] L. Ingber. Very fast simulated re-annealing. *Mathematical and Computer Modelling*, 12:967–973, 1989.

[70] L. Ingber and B. Rosen. Genetic algorithms and very fast simulated re-annealing: A comparison. *Mathematical and Computer Modelling*, 16(11):87–100, 1992.

[71] M. Jenkinson. Measuring transformation error by RMS deviation. Internal Technical Report TR99MJ1, Oxford Centre for Functional Magnetic Resonance Imaging of the Brain, Department of Clinical Neurology, Oxford University, Oxford, UK, 1999. Available at `www.fmrib.ox.ac.uk/analysis/techrep` for downloading.

[72] M. Jenkinson. Improved unwarping of epi volumes using regularised B0 maps. In *Seventh Int. Conf. on Functional Mapping of the Human Brain*, 2001.

[73] M. Jenkinson, P.R. Bannister, J.M. Brady, and S.M. Smith. Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2):825–841, 2002.

[74] M. Jenkinson and S.M. Smith. Global optimisation for robust affine registration. In *Sixth Int. Conf. on Functional Mapping of the Human Brain*, page 479, 2000.

[75] M. Jenkinson and S.M. Smith. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2):143–156, June 2001.

[76] M. Jenkinson, J. Wilson, and P. Jezzard. Perturbation calculation of B0 field for non-conducting materials. In *International Soc. of Magnetic Resonance in Medicine*, 2002.

[77] P. Jezzard. Physiological noise: Strategies for correction. In C.T.W Moonen and P.A. Bandettini, editors, *Functional MRI*. Springer, 1999.

[78] P. Jezzard and R.S. Balaban. Correction for geometric distortion in echo planar images from B0 field variations. *Magnetic Resonance in Medicine*, 34:65–73, 1995.

[79] P. Jezzard and S. Clare. Sources of distrotion in functional MRI data. *Human Brain Mapping*, 8:80–85, 1999.

[80] P. Jezzard, P.M. Matthews, and S.M. Smith. *Functional MRI - An Introduction to Methods*. Oxford University Press, 2001.

[81] A.P. Jiang, D.N. Kennedy, J.R. Baker, R. Weisskoff, R.B.H. Tootell, R.P. Woods, R.R. Benson, K.K. Kwong, T.J. Brady, B.R. Rosen, and J.W. Belliveau. Motion detection and correction in functional mr imaging. *Human Brain Mapping*, 3:224–235, 1995.

[82] H. Johansen-Berg, V. Christensen, M.W. Woolrich, and P.M. Matthews. Attention to touch modulates activity in both primary and secondary somatosensory areas. *Neuroreport*, 11(6):1237–1241, 2000.

[83] G.Q. Maguire Jr, M.E. Noz, H. Rusinek, J. Jaeger, E.L. Kramer, J.J. Sanger, and G. Smith. Graphics applied to medical image registration. *IEEE Computer Graphics and Applications*, 11:20–29, 1991.

[84] L. Junck, J.G. Moen, G.D. Hutchins, M.B. Brown, and D.E. Kuhl. Correlation methods for the centering, rotation and alignment of functional brain images. *Journal of Nuclear Medicine*, 31:1220–1276, 1990.

[85] C. Jutten and J. Herault. Blind separation of sources, part i: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24:1–10, 1991.

[86] B. Kim, J.L. Boes, P.H. Bland, T.L. Chenevert, and C.R. Meyer. Motion correction in fMRI via registration of individual slices into an anatomical volume. *Magnetic Resonance in Medicine*, 41:964–972, 1999.

[87] B. Kim, J.L. Boes, K.A. Frey, and C.R. Meyer. Mutual information for automated multimodal image warping. In *Visualization in Biomedical Computing*, volume 1131 of *Lecture Notes in Computer Science*, pages 349–354. Springer, 1996.

[88] S. Kirkpatrick, C.D. Gelatt Jr., and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, May 1983.

[89] G.J. Klein, B.W. Reutter, and R.H. Huesman. 4D affine registration models for respiratory-gated PET. In *IEEE Nuclear Science Symposium*, 2000.

[90] T. Kochiyama, O. Tomohisa, T. Morita, M. Matsumura, Y. Yonekura, and N. Sadato. Correction and diagnosis of task-related heteroscedasticity in fMRI time-series using independent component analysis. In *Eighth Int. Conf. on Functional Mapping of the Human Brain*, pages 895–896, 2002.

[91] F. Kruggel, D.Y. von Cramon, and X. Descombes. Comparison of filtering methods for fMRI datasets. Max Planck Institute of Cognitive Neuroscience, Leipzig, Germany, August 1998.

[92] J. Kybic, P. Thévenaz, A. Nirkko, and M. Unser. Unwarping of unidirectionally distorted EPI images. *IEEE Transactions on Medical Imaging*, 19(2):80–93, February 2000.

[93] Functional Imaging Laboratory. *Statistical Parametric Mapping: SPM 2.* Wellcome Department of Imaging Neuroscience, May 2003. Available at `http://www.fil.ion.ucl.ac.uk/spm/spm2.html`.

[94] C.C. Lee, R.C. Grimm, A. Manduca, J.P. Felmlee, R.L. Ehman, S.J. Riederer, and C.R. Jack Jr. A prospective approach to correct for inter-image head rotation in FMRI. *Magnetic Resonance in Medicine*, 39:234–243, 1998.

[95] C.C. Lee, C.R. Jack Jr., R.C. Grimm, P.J. Rossman, J.P. Felmlee, R.L. Ehlman, and S.J. Riederer. Real-time adaptive motion correction in functional MRI. *Magnetic Resonance in Medicine*, 36:436–444, 1996.

[96] D. Lemoine, D. Liegeard, E. Lussot, and C. Barillot. Multimodal registration system for the fusion of MRI, CT, MEG, and 3D or stereotactic angiographic data. In *Medical Imaging 1994: Image Capture, Formatting and Display*, volume 2164, pages 46–56. Proc. SPIE, 1994.

[97] A. Lundervold and I.H. Khateeb. Numerical estimation of the BOLD response in event-related fMRI. In *Proc. 25th Annual Mid-Year Meeting of the International Neuropsychological Society*, July 2002.

[98] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximisation of mutual information. *IEEE Trans. on Medical Imaging*, 16(2):187–198, 1997.

[99] J.B.A. Maintz, P.A. van den Elsen, and M.A. Viergever. Comparison of feature-based matching of CT and MR brain images. In N. Ayache, editor, *Computer Vision, Virtual Reality, and Robotics in Medicine*, pages 219–228. Springer, 1995.

[100] J.B.A. Maintz, P.A. van den Elsen, and M.A. Viergever. Comparison of edge-based and ridge-based registration of CT and MR brain images. *Medical Image Analysis*, 1(2):151–161, 1996.

[101] J.B.A. Maintz, P.A. van den Elsen, and M.A. Viergever. Evaluation of ridge-seeking operators for multimodality medical image matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:353–365, 1996.

[102] J.B.A. Maintz and M.A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[103] G. Malandain, S. Fernández-Vidal, and J.M. Rocchisani. Improving registration of 3-D medical images using a mechanical based method. In *3rd European Conference on Computer Vision*, pages 131–136, 1994.

[104] G. Malandain, S. Fernández-Vidal, and J.M. Rocchisani. Rigid registration of 3-D objects by motion analysis. In *12th International Conference on Pattern Recognition*, pages 579–581, 1994.

[105] G. Malandain, S. Fernández-Vidal, and J.M. Rocchisani. Physically based registration of 3-D free-form objects: Application to medical imaging. Technical Report 2453, INRIA Sophia-Antipolis, 1995.

[106] P. Marais, R. Guillemaud, M. Sakuma, A. Zisserman, and J.M. Brady. Visualising cerebral asymmetry. In *Visualisation in Biomedical Computing (LNCS:1131)*, pages 411–416, 1996.

[107] M. J. McKeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski. Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*, 6(3):160–88, 1998.

[108] E.H.W. Meijering, W.J. Niessen, J.P.W. Pluim, and M.A. Viergever. Quantitative comparison of sinc-approximating kernels for medical image interpolation. In *Second International Conference on Medical Image Computing and Computer Assisted Intervention*, volume 1679 of *Lecture Notes in Computer Science*, pages 210–217. Springer, 1999.

[109] M.I. Miller, G.E. Christensen, Y. Amit, and U. Grenander. Mathematical textbook of deformable neuroanatomies. *Proceedings of the National Academy of Science, USA*, 90:11944–11948, 1994.

[110] L. Muresan, R. Renken, J.B.T.M. Roerdink, and H. Duifhuis. Position-history and spin-history artifacts in fMRI time-series. In A. V. Clough and C.-T. Chen, editors, *Proc. Medical Imaging 2002: Physiology and Function from Multidimensional Image*, volume 4683, pages 444–451. Proc. SPIE, 2002.

[111] D.C. Noll, F.E. Boada, and W.F. Eddy. Movement correction in fMRI: The impact of slice profile and slice spacing. *International Soc. of Magnetic Resonance in Medicine*, page 1677, 1997.

[112] J.J. Orchard, C. Greif, G.H. Golub, and M.S. Atkins. Simultaneous registration and activation detection for fMRI. Technical Report SCCM-02-05, Scientific Computing and Computing Mathematics Program, Stanford University, 2002.

[113] J.L. Ostuni, A.K.S. Santha, V.S. Mattay, D.R. Weinberger, R.L. Levin, and J.A. Frank. Analysis of interpolation effects in the reslicing of functional MR images. *Journal of Computer Assisted Tomography*, 21:803–810, 1997.

[114] A. Papoulis. *Probablility, Random Variables, and Stochastic Processes.* McGraw-Hill, second edition, 1984.

[115] A. Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, third edition, 1991.

[116] R.L. Paul. *Robot manipulators.* Series in Artificial Intelligence. The MIT Press, 1981.

[117] C.A. Pelizzari, G.T.Y. Chen, D.R. Spelbring, R.R. Weichselbaum, and C.T. Chen. Accurate three-dimensional registration of CT, PET, and/or MR images of the brain. *Journal of Computer Assisted Tomography*, 13:20–26, 1989.

[118] W. Penny, S. Roberts, and R. Everson. ICA: Model order selection and dynamic source models. In S. Roberts and R. Everson, editors, *Independent Component Analysis: Principle and Practice*, chapter 12. Cambridge University Press, 2001.

[119] J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. Interpolation artefacts in mutual information based image registration. *Computer Vision and Image Understanding*, 77:211–232, 2000.

[120] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipies in C.* Cambridge University Press, second edition, 1995.

[121] R. Prorok and A. Sawyer. *Signa Advantage Applications Guide*, volume 4. GE Medical Systems, 1992.

[122] M.E. Raichle. Visualizing the mind. *Scientific American*, April 1994.

[123] M.D. Robson, J.C. Gatenby, A.W. Anderson, and J.C. Gore. Practical considerations when correcting for movement-related effects present in fMRI time-series. In *International Soc. of Magnetic Resonance in Medicine*, 1997.

[124] R. Rohling, A. Gee, and L. Berman. Three-dimensional spatial compounding of ultrasound images. *Medical Image Analysis*, 1(3):177–193, 1997.

[125] K. Rohr, M. Fornefett, and H. S. Stiehl. Spline-based elastic image registration: integration of landmark errors and orientation attributes. *Computer Vision and Image Understanding*, 90(2):153–168, 2003.

[126] P.J. Rousseeuw and A.M. Leroy. *Robust Regression and Outlier Detection.* Wiley Series in Probability and Mathematical Statistics. Wiley, 1987.

[127] G.A.F. Seber. *Linear Regression Analysis.* Wiley, 1977.

[128] K. Shields, D.C. Barber, and S.B. Sheriff. Image registration for the investigation of atherosclerotic plaque movement. In H.H. Barrett and A.F. Gmitro, editors, *Information Processing in Medical Imaging*, volume 687 of *Lecture Notes in Computer Science*, pages 438–458. Springer, Berlin, 1993.

[129] S.M. Smith, C.F. Beckmann, N. Ramnani, M.W. Woolrich, P.R. Bannister, M. Jenkinson, P.M. Matthews, and D.J. McGonigle. Intersession variability in FMRI and the effect of different analysis methods. In *Ninth Int. Conf. on Functional Mapping of the Human Brain*, 2003.

[130] V.A. Stenger, S. Peltier, F.E. Boada, and D.C. Noll. 3D spiral cardiac/respiratory ordered fMRI data acquisition at 3 Tesla. *Magnetic Resonance in Medicine*, 41:983–991, 1999.

[131] C. Studholme, T.C. Constable, and J.S. Duncan. Non-rigid spin echo MRI registration incorporating an image distortion model: Application to accurate alignment of fMRI to conventional MRI. *IEEE Transactions on Medical Imaging*, 19(11), November 2000.

[132] C. Studholme, D.L.G. Hill, and D.J. Hawkes. Automated 3D registration of MR and CT images of the head. *Medical Image Analysis*, 1(2):163–175, 1996.

[133] C. Studholme, D.L.G. Hill, and D.J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32:71–86, 1999.

[134] J.-P. Thirion. Image matching as a diffusion process: an analogy with Maxwell's demons. *Medical Image Analysis*, 2(3):243–260, 1998.

[135] A.W. Toga and J.C. Mazziotta, editors. *Brain Mapping — The Methods.* Academic Press, San Diego, CA, 1996.

[136] A.W. Toga and P. Thompson. *Brain Warping*, chapter 1. Academic Press, 1999.

[137] C. Tsallis and D.A. Stariolo. Generalized simulated annealing. *Physica A*, 233:395–406, 1996.

[138] P.A. van den Elsen, E.-J.D. Pol, T.S. Sumanaweera, P.F. Hemler, S. Napel, and J.R. Adler. Grey value correlation techniques used for automatic matching of CT and MR brain and spine images. In *Visualization in Biomedical Computing*, volume 2359, pages 227–237. Proc. SPIE, 1994.

[139] A. van der Kouwe and A. Dale. Real-time motion correction using octant navigtors. In *Seventh Int. Conf. on Functional Mapping of the Human Brain*, 2001.

[140] A. Venot, J.L. Golmard, J.F. Lebruchet, L. Pronzato, E. Walter, G. Frij, and J.C. Roucayrol. Digital methods for change detection in medical images. In F Deconinck, editor, *Information Processing in Medical Imaging*, pages 1–16. Nijhoff Publishers, Dordrecht, 1983.

[141] P. Viola and W.M. Wells. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, 1997.

[142] H.A. Ward, S.J. Riederer, R.C. Grimm, R.L. Ehman, J.P. Felmlee, and C.R. Jack Jr. Prospective multiaxial motion correction for fMRI. *Magnetic Resonance in Medicine*, 43:459–469, 2000.

[143] J. West et al. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of Computer Assisted Tomography*, 21(4):554–566, 1997.

[144] C. Westbrook and C. Kraut. *MRI in practice.* Blackwell Science, Oxford, UK, second edition, 1998.

[145] R.P. Woods. Automated Image Registration. Web page, April 1998. `http://www.bishopw.loni.ucla.edu/AIR3/`.

[146] R.P. Woods, S.R. Cherry, and J.C. Mazziotta. Rapid automated algorithm for aligning and reslicing PET images. *Journal of Computer Assisted Tomography*, 16(4):620–633, 1992.

[147] R.P. Woods, S.T. Grafton, J.D.G. Watson, N.L. Sicotte, and J.C. Mazziotta. Automated image registration: II. Intersubject validation of linear and nonlinear models. *Journal of Computer Assisted Tomography*, 22(22):155–165, 1998.

[148] R.P. Woods, J.C. Mazziotta, and S.R. Cherry. MRI-PET registration with automated algorithm. *Journal of Computer Assisted Tomography*, 17(4):536–546, 1993.

[149] M.W. Woolrich, B.D. Ripley, J.M. Brady, and S.M. Smith. Nonparametric estimation of temporal autocorrelation in FMRI. In *Sixth Int. Conf. on Functional Mapping of the Human Brain*, page 610, 2000.

[150] K.J. Worsley and K.J. Friston. Analysis of fMRI time series revisited – again. *NeuroImage*, 2:173–181, 1995.

[151] M.J. Zigmond, F.E. Bloom, S.C. Landis, J.L. Roberts, and L.R. Squire. *Fundamental Neuroscience.* Academic Press, 1998.