ABSTRACT

Title of Dissertation: REGISTRATION METHODS
FOR QUANTITATIVE IMAGING

Gustavo Kunde Rohde, Doctor of Philosophy, 2005

Dissertation directed by: Professor Carlos A. Berenstein and
Professor Dennis M. Healy Jr.
Department of Mathematics

At the core of most image registration problems is determining a spatial transformation that relates the physical coordinates of two or more images. Registration methods have become ubiquitous in many quantitative imaging applications. They represent an essential step for many biomedical and bioengineering applications. For example, image registration is a necessary step for removing motion and distortion related artifacts in serial images, for studying the variation of biological tissue properties, such as shape and composition, across different populations, and many other applications. Here fully automatic intensity based methods for image registration are reviewed within a global energy minimization framework. A linear, shift-invariant, stochastic model for the image formation process is used to describe several important aspects of typical implementations of image registration methods. In particular, we show that due to the stochastic

nature of the image formation process, most methods for automatic image registration produce answers biased towards 'blurred' images. In addition we show how image approximation and interpolation procedures necessary to compute the registered images can have undesirable effects on subsequent quantitative image analysis methods. We describe the exact sources of such artifacts and propose methods through which these can be mitigated. The newly proposed methodology is tested using both simulated and real image data. Case studies using three-dimensional diffusion weighted magnetic resonance images, diffusion tensor images, and two-dimensional optical images are presented. Though the specific examples shown relate exclusively to the fields of biomedical imaging and biomedical engineering, the methods described are general and should be applicable to a wide variety of imaging problems.

# REGISTRATION METHODS
# FOR QUANTITATIVE IMAGING

by

Gustavo Kunde Rohde

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2005

Advisory Committee:

Professor Carlos A. Berenstein and
Professor Dennis M. Healy Jr., Chairman/Advisor
Doctor Peter J. Basser
Professor Ramani Duraiswami
Professor David Jacobs

TABLE OF CONTENTS

# LIST OF FIGURES

# Chapter 1

# Introduction and motivation

Recent advances in digital imaging technology have had profound impact on a variety of technical fields including communications, medicine, surveillance, military, entertainment, as well as many experimental sciences. The availability of charged coupled devices (CCDs), for example, has encouraged widespread use of digital cameras for a variety of purposes ranging from personal entertainment to automated surveillance systems. Microwave-based imaging technologies such as synthetic aperture radar (SAR) are widely used by militaries around the world for the purposes of intelligence gathering. In the biomedical fields, imaging modalities such as magnetic resonance imaging (MRI), computer assisted x-ray tomography (CT), and ultrasound (US), to name a few, are becoming increasingly used for diagnosing, treating, and monitoring pathologies. In addition, many scientific fields such as experimental biology, chemistry, and materials sciences are also becoming increasingly dependent on imaging technologies such as digital atomic force microscopes and high-field magnetic resonance resonance spectroscopy and imaging techniques, to name a few, for acquiring data to be used in validating and even generating scientific hypothesis.

Taking advantage of the ever increasing computational power often cheaply

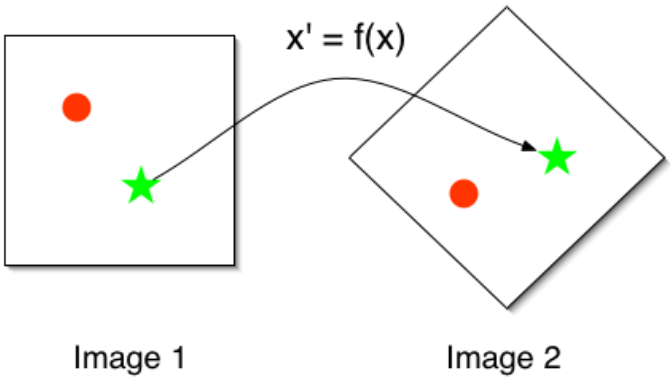


Figure 1.1: Pictorial representation of the image registration process. Because of object or sensor motion, geometric distortions and other artifacts, corresponding objects (circle and star) in two or more instances of the image may not share the same spatial coordinate. Image registration is the process of identifying the spatial correspondence between two or more images.

available, several image processing techniques are currently used to enhance and fuse information from several imaging sources for the purpose of obtaining full benefits from the image data. Paramount for optimal exploitation of available image data is the proper design and application of digital image registration techniques. Image registration refers to the process of identifying the spatial correspondence, as characterized by a spatial transformation of coordinates, between two or more digital images. A pictorial description of this process is provided in Figure (1.1). Precise definitions for the digital images and the spatial transformations that register them will be given later.

Image registration techniques, both manual and automatic, are commonly employed in many of the areas referred to above. In particular, image registration methods have become essential for performing quantitative imaging by combining information from two or more images. Activities such as target recognition and

intelligence gathering can be greatly enhanced by combining information from several sensors placed in different airplanes, satellites, etc. Because of the different position of each sensor, and their different geometric distortion properties, the images first need to be aligned before any meaningful fusion of information is to occur. Automated target tracking, a crucial activity in guided weaponry and robotics, also requires establishing the spatial position of an object, or objects, in a time series of images. Image registration methods can also be used for such purposes.

In the biomedical fields, physicians and research scientists also benefit greatly from image registration methods. It is known that different imaging modalities highlight different properties of biological tissue. X-ray based computed tomography, for example, is known for its ability to provide accurate descriptions of hard tissue types, such as bone, while MRI is know for its sensitivity to different types of soft tissue. Radiologists often benefit from combining CT and MRI modalities for fully characterizing different conditions. Other quantitative imaging modalities such as functional MRI (fMRI), diffusion tensor MRI (DT-MRI), MR relaxometry, etc., depend on fitting models on multiple images acquired using different acquisition parameters. In fMRI researchers are typically interested in mapping the human brain's response to different activities. To that end, multiple (a few tens, or in some cases hundreds) of 3D MR images of a subject's brain are acquired while having the subject perform different activities such as reading, listening, etc. The time series of images is then analyzed by correlating the intensity value of each brain structure with the particular activity. In diffusion tensor MRI the goal is to map the diffusion properties of biological tissues *in vivo* and often in three dimensions. Multiple diffusion weighted MR images are

acquired by introducing diffusion sensitizing magnetic field gradients of different strengths and directions. The intensity values of all the images are then used for fitting a multivariate exponential model, from which it is possible to extract a diffusion tensor that characterizes the amount of local water molecule displacement in three dimensions. Because of subject motion during image acquisition, as well as geometric distortions related to acquisition parameters, the series of acquired images can be severely misaligned with respect to each other. Naturally, before any such automated analysis (fMRI or DT-MRI) are to be extracted from the data, it is absolutely necessary that the multiple images be correctly aligned.

In addition to performing motion and distortion correction in intra-patient data, researchers have also found useful to compare images of the same anatomical part of different subjects (with the human brain being the most prominent example) to study the variation of biological tissue properties, such as shape and composition, described in images across a given population. The aim in such endeavors is usually to provide a quantitative description of a healthy 'normal' population, as opposed to a diseased one. Though a significant amount of research has been performed in the area of inter-subject comparisons, the area is relatively new, and results of substantial impact are lacking. Moreover, citing concerns over homology and correspondence, many researchers have expressed well founded skepticism of many methods and results obtained with them [23, 28].

Because of its crucial importance for many imaging applications image registration methods have been extensively studied. For recent reviews refer to [72] and [126, 84]. Because the application areas greatly vary, image registration methods have been studied by numerous scientists with varied scientific backgrounds, including electrical engineers, applied mathematicians, scientific computation ex-

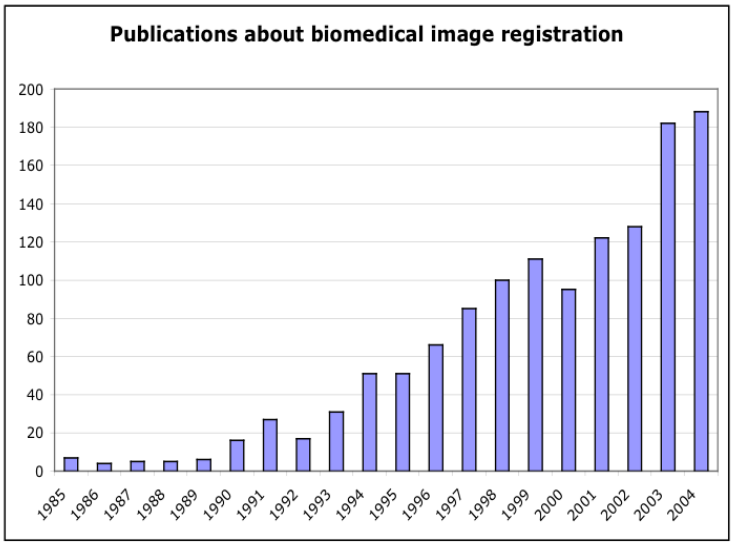


Figure 1.2: Number of peer-reviewed journal publications about biomedical image registration per year. Though this represents a lower bound estimate, the trend exposes the increasing impact of image registration methods on the clinical and biomedical communities.

perts, physicians, physicists, chemists, biologists, psychiatrists, statisticians and possibly others. A quick study using the PubMed [80] search engine yields an estimate of the number of peer reviewed journal publications about biomedical image registration per year. The search term used was 'image AND registration.' The trend is shown in Figure (1.2). Of course, this is a lower bound estimate since this particular search engine specializes in biomedical and clinical applications, and may exclude some more technical journals. Moreover, the number of publications about image registration in areas other than clinical or biomedical is likely to increase the totals by a significant amount. Nonetheless, the figure is useful to expose the trend of increasing importance image registration methods have had on the clinical and biomedical communities over the past two decades.

## 1.1 Problem statement

As explained above, post-acquisition image alignment (registration) is routinely performed in biomedical research and clinical practice [72, 84]. Applications using image registration techniques include motion and distortion correction in fMRI, DT-MRI, and MR relaxometry experiments. In addition, image registration procedures are increasingly being used in computational based studies of neuroanatomy. This involves understanding the variability of tissue properties, including shape, across specific populations. An example is voxel-based morphometry, described in [11].

In general, many of the current post-processing methodologies can be summarized within a pipeline framework, as depicted in Figure (1.3). At first, a set of medical images is acquired and reconstructed using standard tomographic technologies. The tomographic reconstruction step in MRI typically involves a Fourier transformation of the data (though filtered back-projection methods are sometimes used) while reconstruction procedures in CT often involve Radon transformation methods. Regardless of the tomographic reconstruction method in use, in most quantitative imaging experiments the output of this step is a series of digital images to be stored in computer memory. Each image in this series can be though of as a function (real or complex) of discrete input coordinates. Mathematically, the $n^{\text{th}}$ image in this series is written as $S_n(\mathbf{i})$, with $\mathbf{i} \in \mathbb{Z}^d$, where $d$ represents the dimensionality of the images. For example, a common digital image processing technique is to view the indexes of the image array $\mathbf{i}$ as $\mathbf{i} = [i, j, k]^T$ where each coordinate $i, j, k$ belongs to the set $\{0, \cdots, 255\}$, for example. Naturally, the coordinates $\mathbf{i}$ are associated with the spatial coordinate system of the laboratory. The specifics of this association are determined by the

image acquisition and reconstruction procedure, but simplifications can be made such that a coordinate in the laboratory frame of view $\mathbf{x} \in \Omega := [0,1]^d \subset \mathbb{R}^d$ can be written as $\mathbf{x} = [c_x i, c_y j, c_z k]^T$. The constants $c_x, c_y, c_z$ represent the resolution (size of each sample) in the $x, y$ and $z$ dimensions, respectively, of the imaging system.

Because of patient motion, or device dependent geometric distortions which may not remain constant through acquisition of the entire image series, the series of images $S_n(\mathbf{i})$, $1 \leqslant n \leqslant N$, may be misaligned with respect to each other. This means that a fixed image coordinate $\mathbf{i}$ may not represent the same structure or anatomical region in all images of the series. To ensure that the same coordinate $\mathbf{i}$ corresponds, as much as possible, to the same structure image registration is performed to bring the series of images into alignment (see figure (1.3)). This entails in finding functions $f_n : \Omega \to \Omega$, $1 \leqslant n \leqslant N$, so that the variability due to subject motion or geometric distortions in the series $S_n(f_n(\mathbf{x}))$, $1 \leqslant n \leqslant N$, is removed. In addition to removing artifacts related to motion and distortion the entire image series may also be spatially transformed onto a standardized coordinate system so that the data may be more conveniently interpreted. More on how these tasks are actually accomplished is to follow.

Once variability due to patient motion and geometric distortions have been accounted for, the next step in a typical image processing pipeline is to use the registered data for some quantitative analysis. Data analysis consists of extracting or estimating some physically meaningful parameters from the sequence of image data. Typically, this is done by fixing a coordinate $\mathbf{x}$ and defining a data vector $\mathbf{y} = [S_1(f(\mathbf{x})), \cdots, S_N(f(\mathbf{x}))]^T$. Next a model, whose quantitative parameters may contain useful physical information, is extracted from the data $\mathbf{y}$

usually through estimation procedures such as the maximum likelihood method. In DT-MRI, for example, the model parameters define a 3x3 symmetric effective diffusion tensor. In fMRI and voxel based morphometry, the parameters define statistical parametric maps. For such analysis to be meaningful, the vector $\mathbf{y}$ must consist of image data values from the same anatomical region. For example, if one wishes to analyze the effective diffusion properties of the human brain's corpus callosum it is paramount that the data $\mathbf{y}$ contain image values only from the corpus callosum. Image misalignment may introduce data values from other sources such as air (background) or cerebral spinal fluid in $\mathbf{y}$, causing significant errors in the estimated parameters. Thus, a large portion of quantitative imaging applications, especially in biomedicine, depend on accurate image registration methods.

## 1.2 Summary of Dissertation

Despite intense scrutiny and study, many unresolved issues regarding image registration methodology remain. Many of the difficulties currently encountered can be attributed to the fact that image processing specialists tend to formulate the problem without much consideration for the physical properties of the image formation process. Likewise, scientists interested in applications such as fMRI and DT-MRI tend to use image registration (and other image processing methods) as a black box and usually exclude details about pre-processing steps in their data analysis. Such disconnect is understandable given that a successful quantitative imaging experiment involves scientists with a variety of backgrounds: from imaging physicists and engineers at the data acquisition end, to biologists and physicians concerned in making a specific discovery at the output end, and with

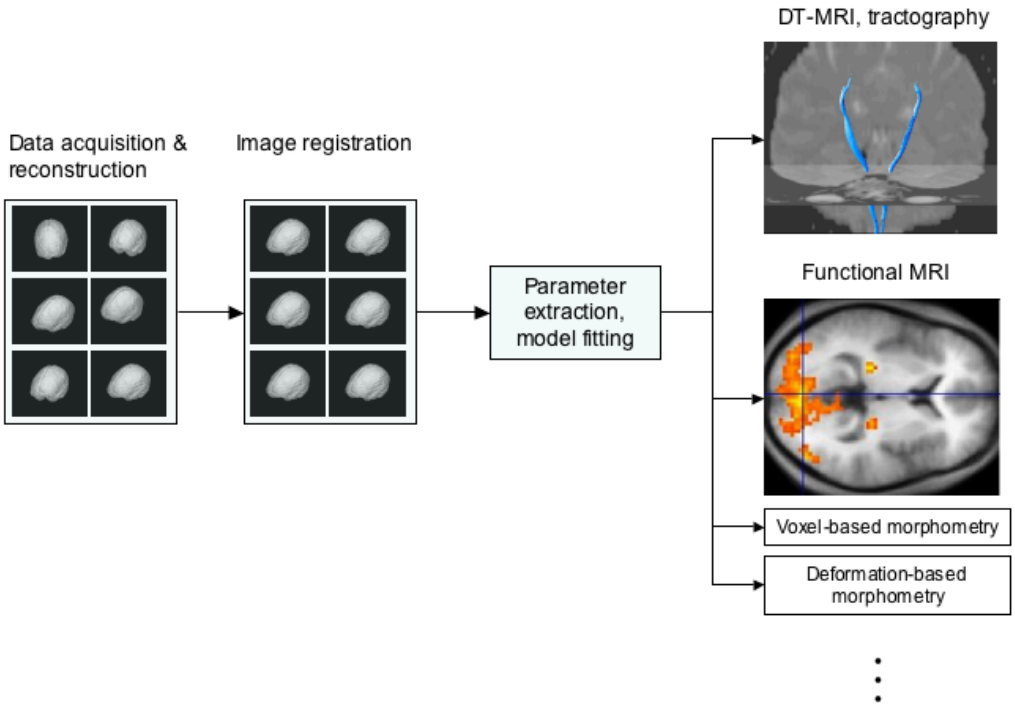


Figure 1.3: Overview of a typical biomedical quantitative image processing pipeline. First the images are acquired and reconstructed using one of the many available tomographic reconstruction techniques. Next the images are registered (aligned) to ensure, as much as possible, that a fixed image coordinate corresponds to the same anatomical structure in all images. Finally, the registered data are used to produce quantitative parameters through model fitting. The output of the pipeline is then used by researchers in the biomedical fields for testing and generating hypotheses, as well as for diagnostic and therapeutic applications.

applied mathematicians and scientific computation experts in the middle being responsible for several data processing steps.

This work addresses some of the shortcomings of current image registration, and data processing methods in general, by incorporating, as much as possible, information about the image formation process into the formulation of the image registration problem. Improvements over currently available technology are described not only for finding the solution of the registration problems but also in using the results for subsequent quantitative image analysis. Though all of the applications investigated in detail here are of biomedical nature, the ideas to be presented are general and could be used in many technical fields that depend on image data.

Next, some of the main ideas used in biomedical image registration throughout the years are reviewed so that the original contributions presented in this work can be more easily identified and understood. Chapter 3 provides a general overview of the original contributions to be presented in subsequent chapters. The image registration problem is analyzed using a linear, shift-invariant, stochastic image model. Equations and numerical methods to address the image registration problem are derived in general, without specific attention to any single application. In particular, we analyze the effects of object (sensor) motion, geometric distortions, and thermal noise in serial images and propose methods through which such artifacts can be addressed.

Chapter 4 investigates approaches for measuring the similarity between two images, paying particular attention to thermal noise, in the context of solving the image registration problem. In it, we show that the vast majority of image registration methods implemented, by default, are not likely to produce optimal

answers. Rather, the solution achieved with current methods is likely to be biased towards the most 'blurred' image. Methods through which such shortcoming can be addressed are presented.

In chapter 5 the methodology described earlier is used to address the problem of patient motion and eddy-current induced image distortions in serial acquisitions of diffusion weighted MRI. In particular, the linear, shift-invariant, modeling approach is used to derive appropriate spatial transformation models to be used in the solution of the problem. Results presented using real patient data show a significant improvement in the quality of the diffusion studies obtained. Another case study is presented in chapter 6, where we investigate motion artifacts induced by muscle contraction in epicardial fluorescence optical imaging studies. Again results obtained from real data experiments show that the motion artifacts can be largely minimized using image registration methods.

Finally, in chapter 7 the effects of image interpolation and resampling methods on quantitative image applications such as diffusion tensor MRI are investigated. In particular, the effects of image intensity variance due to noise are propagated through the image registration process to that maximum likelihood and least squares fitting procedures, for example, can be properly implemented for extracting quantitative model parameters. Concluding remarks are offered in chapter 8.

# Chapter 2

# Background

Here we provide a brief description of some of the main ideas researchers have used to address several aspects of the registration problem defined in the previous chapter. The following discussion is not meant to represent an exhaustive description of the currently available literature about image registration (for that the reader is referred to [72, 126, 84, 29, 45, 76]), but rather to clarify and relate the original ideas presented herein as compared to other work. Though image registration methods can be categorized according to many criteria (see [72] for good examples) we divide them into two main categories: 1) feature-based methods which rely on substantial amounts of, sometimes manual, pre-processing and 2) fully automatic intensity based methods.

## 2.1 Feature-based methods

Feature-based registration methods depend on the identification of corresponding landmarks, curves or surfaces, for computing the spatial transformation that brings two images into spatial alignment. Let $T(\mathbf{x}), \mathbf{x} \in \Omega := [0,1]^d \subset \mathbb{R}^d$ represent a reference or target image to which we would like to align a source image $S(\mathbf{x}), \mathbf{x} \in \Omega$. Let $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_M]$ represent a series of $M$ points $\mathbf{x}_n \in \Omega$ that

correspond to a series of landmark points in the target image $T(\mathbf{x})$. Suppose that, somehow, the coordinates of the corresponding structures in image $S(\mathbf{x})$ are known. In the case of head image registration, such landmarks can be obtained, manually or automatically, by using a fiducial object such as a stereotactic frame screwed rigidly to the patient's outer skull [68]. Methods for identifying significant landmark points based on intensity values alone also exist. Corresponding landmarks can be identified manually [34] or automatically [48]. Let the corresponding coordinates in image $S(\mathbf{x})$ be denoted $\mathbf{Y}$. The coordinates in the source image $\mathbf{Y}$ can be written, approximately or sometimes exactly, as a function of the target image coordinates $U(\mathbf{X}, \mathbf{c})$, where $\mathbf{c}$ represents the vector of parameters that defines the spatial transformation. This can be done by minimizing the error between the coordinate points in a least-squares sense:

$$\min_{c} \|\mathbf{Y} - U(\mathbf{X}, \mathbf{c})\| \tag{2.1}$$

When the spatial transformation $U(\cdot)$ is a rigid body one, i.e. $U(\mathbf{X}) = \mathbf{R}\mathbf{X} + \mathbf{t}$, where $\mathbf{R}$ is an orthogonal rotation matrix and $\mathbf{t}$ is a translation vector, the solution can be computed analytically in a least squares sense [97]. This approach has been used extensively for registering images of different modalities [33, 52].

More generally, the transformation $U(\cdot)$ may be an elastic one in which points $\mathbf{X}$ are displaced non-uniformly so as to match the set of points $\mathbf{Y}$. A technique often used is to model $U(\cdot)$ as a linear combination of symmetric, radial basis functions $\phi(\cdot)$ as:

$$U(\mathbf{x}, \mathbf{c}) = \sum_{i=1}^{M} c_i \phi(\|\mathbf{x} - \mathbf{x}_i\|), \tag{2.2}$$

where $\mathbf{x}_i$, $1 \leqslant i \leqslant M$, are the landmark points in the target image space, and, $c_i$

are the coefficients of the expansion. If the basis functions are positive definite in the sense that the matrix $\mathbf{M}_{j,k} = \phi(\|\mathbf{x}_j - \mathbf{x}_k\|)$ is positive definite, then the coefficients $c_i$ in (2.2) can be solved for so that the landmark points in both images match exactly. That is $\mathbf{y}_n = U(\mathbf{x}_n, \mathbf{c})$ for all landmark points $n$. The procedure outlined here is a specific example of data interpolation and approximation in multiple dimensions. More details about currently available techniques can be found in [96, 22]

Once the transformation $U(\mathbf{x}, \mathbf{c})$ has been defined, any point $\mathbf{x}$ in the domain of the target image can be transfered to the source image domain. A simple image interpolation or approximation procedure can then be used to resample the source image so as to match the target image. The match should be almost exact at the landmark points, the limitation being how well corresponding landmark points can be defined. A not very flattering illustrative example of this procedure is provided in figure (2.1), where the two-dimensional thin-plate spline basis function was used to define the transformation in (2.2):

$$\phi(r) = -r^2 \log(r^2). \tag{2.3}$$

The source image was resampled using bilinear interpolation (two-dimensional linear interpolation).

The procedure outlined above can be expanded for matching curves in 2D, or surfaces in 3D. The idea is very simple, and it is based on using the segmented curves or surfaces as the sole input for computing the registration. The segmentation can be performed manually. However, because this can be a labour intensive procedure, especially in 3D, automatic segmentation methods based on deformable models and level set methods are often used [125, 120]. Once properly segmented, the anatomical elements are matched with their counterparts

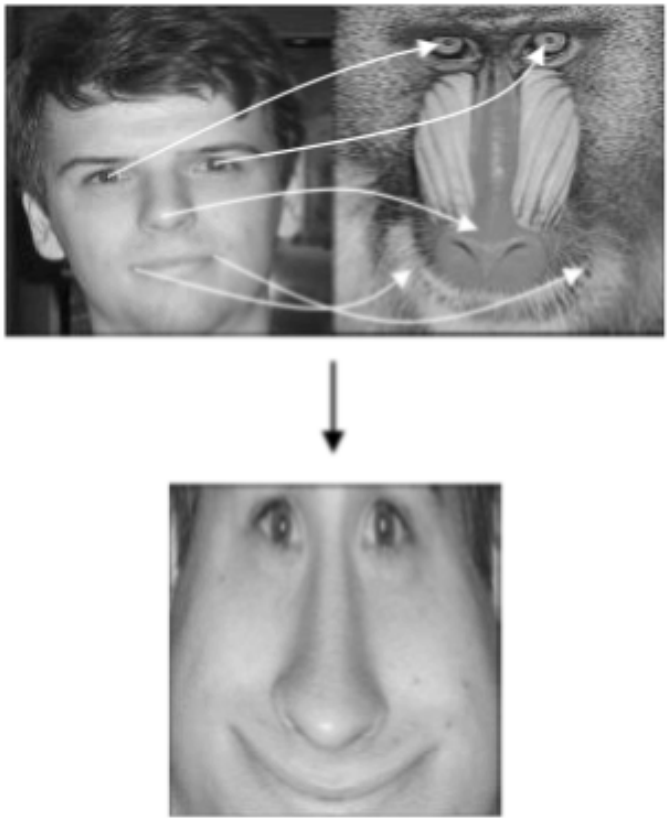

Figure 2.1: Manual, landmark-based image registration example. A set of points $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_5]$ was manually selected in the baboon image. A set of corresponding points $\mathbf{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_5]$ was selected in the human image. The transformation $\mathbf{y}_n = U(\mathbf{x}_n, \mathbf{c})$, as defined by equation (2.2) was computed according to the procedure outlined above. The source (human) image was resampled using bilinear interpolation. The overall effect is to warp the source (human) image so that points $\mathbf{Y}$ "move" to the locations $\mathbf{X}$, as indicated by the white arrows. Note that the intensity values of the human image are not significantly changed. Rather it is the the position in space of each structure that is modified.

in other images [82, 110, 30, 74]. Their correspondence is then used to guide the two-dimensional or volumetric transformation from one image to another. Naturally, the computational problem becomes nonlinear and algorithms become iterative.

There are several drawbacks associated with the procedures discussed above. First, the accuracy of the methods is limited to the initial feature extraction procedure (segmentation of landmarks, curves and surfaces). Though significant research in automatic segmentation methods has been performed, current technology is still susceptible to errors caused by local optima artifacts. Methods for deformable model segmentation, for example, are known to be extremely dependent on initialization. Lastly, such methods often provide a relatively sparse match. That is, even though correspondence between general features can be obtained, the correspondence between remaining structures is not defined. For example, even though the correspondence between two surfaces can be computed, the correspondence of structures inside or outside the surface is not determined. For these and other reasons, the focus of much research related to image registration has shifted from feature-based registration methods to more 'dense' registration methods that use most or all available information in the images, with little or no pre-processing, to compute the spatial correspondence between two or more images. Such methods are termed intensity-based registration methods and are discussed next.

## 2.2 Intensity-based methods

Intensity-based image registration methods differ from the approaches described above in the sense that they generally do not depend on prior feature extraction or

pre-processing of the image data, but rather use all the available intensity value information in order to compute the transformations that register the images. Through interpolation and approximation techniques (discussed in more detail later) the images are viewed as a continuum, and spatial transformations are sought such that the images being registered become more similar:

$$S(f(\mathbf{x})) \sim T(\mathbf{x}). \tag{2.4}$$

An overview of the principal ideas commonly used in intensity-based image registration methods is provided here. We start by defining the image registration problem within a global variational energy minimization context. The discussion is followed by more detailed descriptions of algorithms and methodologies in current practice. The main ideas used in modeling of deformation fields, design of image similarity measures, numerical optimization strategies, and image interpolation/approximation methods for the problem of image registration are discussed.

### 2.2.1 Problem statement

Let $S(\mathbf{x})$ and $T(\mathbf{x})$ be real valued continuous functions on a bounded domain $\Omega := [0,1]^d$, say the unit square or unit cube. The goal in image registration is to obtain a continuous, differentiable, function $f : \Omega \to \Omega$, with continuous and differentiable inverse (diffeomorphism) that minimizes a functional (a non-negative scalar measure on $f$):

$$\min_f \Theta(f(\mathbf{x})). \tag{2.5}$$

The functional $\Theta(\cdot)$ is meant to serve as a measure of image dissimilarity. For simplicity, one may think of $\Theta(\cdot)$ a being a functional of input images $T(\mathbf{x})$ and $S(f(\mathbf{x}))$:

$$\Theta(f(\mathbf{x})) = \int_{\Omega} \Phi(T(\mathbf{x}), S(f(\mathbf{x})))d\mathbf{x}. \tag{2.6}$$

In the simplest cases, $\Phi(y_1, y_2) = (y_1 - y_2)^2$. It is worthwhile pointing out that the problem defined by equations (2.5,2.6) is ill posed in the sense that there may exist many distinct functions $f$ indistinguishable from each other under functional (2.6). One obvious trivial example would be when both images $S(\mathbf{x})$ and $T(\mathbf{x})$ are equal to a constant, say 1, for some connected set of coordinates in $\Omega$. In this case, the common solution is to restrict the search space and give more preference to functions $f$ which contain some smoothness properties [111]. This can be done explicitly through parameterization of $f$, say, modeling $f$ as an affine transformation as explained in a following section, or by adding a penalty term to (2.6) such that:

$$\Theta(f(\mathbf{x})) = I(T(\mathbf{x}), S(f(\mathbf{x}))) + C(f) \tag{2.7}$$

where $I(\cdot, \cdot)$ is the image dissimilarity measure (the negative of the similarity measure) and $C(f)$ is the constraint term. In many applications, $I(\cdot, \cdot)$ is given by equation (2.6), while the constraint term is usually measured as the energy of some differential operator on $f$:

$$C(f) = \int_{\Omega} \Psi(f(\mathbf{x}))d\mathbf{x}, \tag{2.8}$$

where $\Psi(\cdot)$ usually represents a differential operator.

Many methods try to solve the image registration problem by parameterizing $f$, say with rotations or translations (see next section), and explicitly minimizing the equations above. Other examples can be derived by using standard techniques from variational calculus. Let $f(\mathbf{x})$ be a minimizer for equation (2.7) which satisfies some known boundary conditions $f(\mathbf{x}) = b(\mathbf{x}), \forall \mathbf{x} \in d\Omega$. Choose any differentiable function $v(\mathbf{x})$, with $v(\mathbf{x}) = 0 \forall \mathbf{x} \in d\Omega$, and substitute $f(\mathbf{x})$ in (2.7) with $f(\mathbf{x}) + \epsilon v(\mathbf{x})$. Equation (2.7) can thus be written as a function of $\epsilon$:

$$\Theta(\epsilon) = \int_{\Omega} F(\mathbf{x}, f + \epsilon v, f^{'} + \epsilon v^{'}) d\mathbf{x}, \tag{2.9}$$

where $\int_{\Omega} F(\cdots) d\mathbf{x} = \int_{\Omega} \Phi(\cdot, \cdot) + \Psi(\cdot) d\mathbf{x}$, as defined above. Note that in general higher order derivatives may be included. If $f(\mathbf{x})$ is indeed a minimizer, we must have:

$$\frac{d\Theta(\epsilon)}{d\epsilon} = 0. \tag{2.10}$$

It can be shown that $\frac{d\Theta(\epsilon)}{d\epsilon}$=0, as $\epsilon \to 0$, is equivalent to:

$$\int_{\Omega} \left[ \frac{\partial F}{\partial f} v + \frac{\partial F}{\partial f'} \frac{dv}{d\mathbf{x}} \right] d\mathbf{x} = 0. \tag{2.11}$$

Note that the partials $\frac{\partial F}{\partial f}, \frac{\partial F}{\partial f'}$ are formed by treating $\mathbf{x}, f$ and $f^{'}$ as independent variables. This equation is known as the weak form of the variational energy and can be used in finite element-based solutions to the registration problem. Noting that $v(\mathbf{x})$ vanishes at the boundary, integration by parts can be used to show that the weak form above simplifies to:

$$\int_{\Omega} \left[ \frac{\partial F}{\partial f} - \frac{d}{d\mathbf{x}} \frac{\partial F}{\partial f'} \right] v d\mathbf{x} = 0. \tag{2.12}$$

If $f(\mathbf{x})$ is truly a minimizer, the equation above must be valid for all $v$ that vanish at the boundary, as explained earlier. Therefore, if $f(\mathbf{x})$ minimizes the integral equation 2.7, then it must satisfy the Euler-Lagrange equation defined as:

$$\frac{\partial F}{\partial f} - \frac{d}{d\mathbf{x}}\frac{\partial F}{\partial f'} = 0. \tag{2.13}$$

More technical aspects of the variational formalism used can be found in [51]. This equation can be used to derive many finite difference-based image registration schemes.

Whether consciously or not, the above optimization framework is used by numerous researchers aiming to compute the registration between two or more images. In section 2.2.5 we provide a few specific examples in which the equations above have been used to solve typical image registration problems. In the next section we discuss common approaches through which modeling of the spatial transformations $f(\mathbf{x})$ that register the images is performed.

### 2.2.2 Spatial transformations

Depending on the specific application, equation (2.7) can be minimized using many different kinds of spatial transformations. In the simplest examples, rigid body transformations are used. The transformation takes the form:

$$f(\mathbf{x}) = \mathbf{R}\mathbf{x} + \mathbf{t}, \tag{2.14}$$

where $\mathbf{R}$ is a pure rotation matrix, and $\mathbf{t}$ represents a translation vector. While this is an appropriate model in many applications, examples include motion correction in serial images of the head and brain [123], it lacks flexibility to address many other important problems in medical imaging. The next step is to add

affine parameters to the transformation in (2.14). That is, the rotation matrix $\mathbf{R}$ is replaced by an affine transformation $\mathbf{A}$ which not only includes rotations, but also scalings and shears. Affine transformations have proved useful for gross-scale comparisons of brain images of different subjects [70, 11]. In addition, affine transformations are frequently used to remove linear geometric distortions in serial MR images [47, 73]. Moreover, optimization, even in the 3D case is not prohibitively computationally expensive due to the fact that only a few parameters (12 in the case of affine) need to be optimized.

Many imaging applications, such as tracking deformable tissues (examples include liver [94] and heart [98] registration) over time, require nonlinear image coordinate displacements. The next logical step is to include polynomial terms in the spatial transformation $f(\mathbf{x})$ [124]. For example, let $f(\mathbf{x}) = \{x^{'}, y^{'}, z^{'}\}^T$:

$$x^{'} = P_{x0} + P_{x1}x + P_{x2}y + P_{x3}z + P_{x4}x^2 + P_{x5}xy + P_{x6}xz + P_{x7}y^2 + P_{x8}yz + P_{x9}z^2$$
$$y^{'} = P_{y0} + P_{y1}x + P_{y2}y + P_{y3}z + P_{y4}x^2 + P_{y5}xy + P_{y6}xz + P_{y7}y^2 + P_{y8}yz + P_{y9}z^2$$
$$z^{'} = P_{z0} + P_{z1}x + P_{z2}y + P_{z3}z + P_{z4}x^2 + P_{z5}xy + P_{z6}xz + P_{z7}y^2 + P_{z8}yz + P_{z9}z^2.$$

Naturally, the number of parameters to optimize ($P_{.,.}$) increases substantially as the order of the model increases. A similar idea is used by Ashburner and Friston [10]. Instead of polynomials, however, the spatial transformation is composed of cosine basis functions, parameterized by their coefficients, and distributed over a regular grid in domain $\Omega$:

$$f(\mathbf{x}) = \mathbf{x} + \sum_{\mathbf{i} \in \mathbb{Z}^d} \mathbf{c}_\mathbf{i} \phi_\mathbf{i}(\mathbf{x}), \tag{2.15}$$

where $\mathbf{c} \in \mathbb{R}^d$ is a vector of basis function coefficients, with $d$ being the dimensionality of the images. The basis functions $\phi_i(\mathbf{x})$ are separable in the sense:

$$\phi_i(\mathbf{x}) = \prod_{j=1}^{d} \phi_i(x_j), \tag{2.16}$$

while the 1D basis functions are given by the discrete cosine transform of an $N$ dimensional discrete signal [10]. The model above can be expanded to any number $N \times d$ of degrees of freedom. Typically the authors choose to model the deformation field with an $8 \times 8 \times 8$ grid of 3D basis functions, resulting in about 1500 parameters to optimize.

Because of the large number of optimization parameters combined with the fact that cosine basis functions do not have compact support, the spatial transformation model described above may take a long time to compute. This is especially true if the similarity measure being optimized is itself computationally expensive. To alleviate this problem Rueckert and colleagues [95] opt for using compactly supported b-splines for modeling the deformation field expressed in (2.15). Thus, in their implementation $\phi_i(x) = \beta^3(x - x_i)$ with $x_i$ representing the center coordinate for the $i^{\text{th}}$ basis function and

$$\beta^3(x) = \begin{cases} \frac{2}{3} - \frac{1}{2}|x|^2(2 - |x|) & , 0 \leq |x| < 1; \\ \frac{1}{6}(2 - |x|)^3 & , 1 \leq |x| < 2 \ ; \\ 0 & , 2 \leq |x| \ . \end{cases} \tag{2.17}$$

Realizing that nonlinear transformations with a large number of degrees of freedom may caused undesirable artifacts such as 'folding' or 'tearing' of the image, the authors choose to constrain the optimization so as they be as smooth as possible by defining the constraint term (2.8) with

$$C(f) = \int_{\Omega} \| \frac{d^2 f(\mathbf{x})}{d\mathbf{x}^2} \|^2 d\mathbf{x}. \tag{2.18}$$

Both deformation field modeling methods described above (cosine and b-splines) are based on a linear combination of basis functions organized on a regular grid. In some situations very fine grids of basis functions are needed to capture the necessary local deformations. One such application is inter-subject brain comparisons where the cortical variability between different subjects may be high. Such modeling approach may lead to optimization of hundreds of thousands of basis functions. In oder to reduce the complexity of the spatial transformation Rohde *et al.* [87] propose an adaptive modeling scheme based on compactly supported, radially symmetric basis functions. The idea is to concentrate degrees of freedom in regions where optimization is most likely to be successful while minimizing computation in other parts of domain $\Omega$. The determinant of the Jacobian matrix of the computed deformation field $f(\mathbf{x})$ is constrained to remain positive via a computationally efficient inequality derived based on operator theory arguments.

Finally, there is a class of deformation fields that are non-parametric in the sense that an explicit model (basis functions, affine, etc.) for the spatial transformation is not assumed. Rather, the deformation field is estimated using diffeomorphic formulations derived based on assumptions from continuum mechanics. These deformation fields are best described through the constraint function $C(f(\mathbf{x}))$ being used. Linear elasticity models are often used for constraining deformation fields. These can normally be expressed as some (spatial) differential operation on the deformation field $\mathbf{u} = f(\mathbf{x})$. The energy of second derivatives of $f(\mathbf{x})$, such as the one in (2.18), are often used.

Fluid dynamics-type regularizations are also used. The key idea in these methods is not to impose smoothness constraints on the deformation field $\mathbf{u}$ itself,

but rather on its rate of change. Thus, an 'artificial' time variable is introduced and used in the iterative solution of the registration problem. A good example is found in the work of Christensen *et al.* [26] who use the Cauchy-Navier operator in conjunction with the first variation of the 2-norm of the difference image. The governing equations of the method are:

$$\mathbf{b} = \mu\nabla^2\mathbf{v} + (\lambda + \mu)\nabla(\nabla \cdot \mathbf{v}), \tag{2.19}$$

where the material properties are determined by the constants $\lambda$ and $\mu$, and $\mathbf{b}$ represents the first variation of the 2-norm of the difference image. The rate of change of the deformation is computed as:

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{v} - \mathbf{v} \cdot \nabla\mathbf{u}. \tag{2.20}$$

Simple Euler time step integration is used to compute $\mathbf{u}$ at each step based on the equations above.

Finally, regularizations based on the determinant of the Jacobian matrix $|J(f(\mathbf{x}))|$

$$C(f(\mathbf{x})) = \int_\Omega \varphi(|J(f(\mathbf{x}))|)d\mathbf{x} \tag{2.21}$$

are also used [87].

The regularization approaches described above are typically used in PDE-based implementations derived based on the global variational energy minimization framework described above. Some of these will be described in detail in section 2.2.5. Lastly, note that the regularization functions described above can, and are, also used in parametric optimization approaches.

### 2.2.3 Image dissimilarity measures

Once a model for the spatial transformation $f(\mathbf{x})$ is chosen, the image dissimilarity $I(\cdot,\cdot)$ part of equation (2.7) must be defined. The choice of $I$ is usually determined by the requirements of the application. In instances when images $S$ and $T$ are expected to be nearly identical under optimal alignment, such as when matching images from the same image sensor, $L_2$-norms or the sum of squared differences are often employed:

$$I(T(\mathbf{x}), S(f(\mathbf{x}))) = \int_{\Omega} |T(\mathbf{x}) - S(f(\mathbf{x}))|^2 d\mathbf{x}. \tag{2.22}$$

The image $S(f(\mathbf{x}))$ is usually computed using multidimensional linear interpolation, though other options certainly exist and will be discussed in a later section. The equation above may be the most frequent in computer implementations of image registration methods due to its low computational requirements, easy differentiation with respect to the parameters that define spatial transformation $f(\mathbf{x})$, and ease of implementation. Note also that in practice, equation (2.22) is implemented as:

$$I(T(\mathbf{x}), S(f(\mathbf{x}))) \cong \sum_{i=1}^{N} (T(\mathbf{x}_i) - S(f(\mathbf{x}_i)))^2, \tag{2.23}$$

where $N$ is the total number of sampling coordinates $\mathbf{x}_i$ chosen.

In many applications the constant image intensity assumption is not appropriate. When the relationship between the intensity values of $S$ and $T$ is unknown but expected to be linear (such as when matching images from the same type of sensor) $I$ can be based on the linear cross correlation (zero mean correlation coefficient) of their intensity values:

$$I(T(\mathbf{x}), S(f(\mathbf{x}))) = \frac{\langle T(\mathbf{x}), S(f(\mathbf{x}))\rangle}{\|T(\mathbf{x})\|\|S(f(\mathbf{x}))\|}, \tag{2.24}$$

where

$$\|S(\mathbf{x})\|^2 = \langle S(\mathbf{x}), S(\mathbf{x})\rangle = \int_\Omega |S(\mathbf{x})|^2 d\mathbf{x}. \tag{2.25}$$

Again, in practice the integral defined above is computed using discrete samples $\mathbf{x}_i$, $1 \leqslant i \leqslant N$ organized on a regular grid:

$$\langle S(\mathbf{x}), S(\mathbf{x})\rangle \cong \frac{1}{N}\sum_{i=1}^{N} |S(\mathbf{x}_i)|^2. \tag{2.26}$$

Though heavily used, the image similarity measures described above are not appropriate for matching images of different modalities. Consider, for example, the task of finding the spatial correspondence between MR and CT images. Because the physics of the imaging process of MR and CT images differ substantially, so do their intensity values. CT, for example, is known for providing high intensity values for bony structures, while the intensity value due to bone is low in MR. Because of the complex nature of the problem a definitive model that relates the intensity values in these two image modalities has yet to come forth. To circumvent such difficulties during image registration researchers have adopted more general measures of statistical dependency between the images' intensity values as a way of measuring their similarity. The primary example of such approaches involves the use of the mutual information similarity measure [61], first used in medical image registration problems in [56, 70].

Let $T$ represent a random variable associated with the experiment of choosing an intensity value from image $T(\mathbf{x})$ at random positions $\mathbf{x}$ chosen from the entire domain of the images. The probability density function (*pdf*) of $T$ is denoted

$P_T(t)$, where $t$ refers to the probability of outcome of the experiment. Similarly let $S_f$ be the random variable associated with the experiment of choosing, randomly, an intensity value from image $S(f(\mathbf{x}))$. Similarly, we denote the *pdf* of this experiment as $P_{S_f}(s)$. Finally, let $P_{S_f,T}(s,t)$ represent the (joint) *pdf* of random variables $S_f$ and $T$. The mutual information of random variables $T$ and $S_f$ is defined as:

$$I\left(S_f, T\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P_{S_f,T}(s,t)\log\left(\frac{P_{S_f,T}(s,t)}{P_{S_f}(s)P_T(t)}\right)dsdt, \tag{2.27}$$

where the integrals, unless noted otherwise, are taken from $-\infty$ to $\infty$. Alternatively, the mutual information can also be expressed in terms of the entropy of the random variables:

$$I\left(S_f, T\right) = H\left(T\right) + H\left(S_f\right) - H\left(S_f, T\right), \tag{2.28}$$

with

$$H\left(T\right) = -\int P_T(t)\log\left(P_T(t)\right)dt, \tag{2.29}$$

$$H\left(S_f\right) = -\int P_{S_f}(s)\log\left(P_{S_f}(s)\right)ds, \tag{2.30}$$

and

$$H\left(S_f, T\right) = -\int\int P_{S_f,T}(s,t)\log\left(P_{S_f,T}(s,t)\right)dsdt. \tag{2.31}$$

Naturally, the *pdf*'s needed to perform the computations above are not generally available. Rather, all that is available are (sampled) instances of the images $T(\mathbf{x})$ and $S(f(\mathbf{x}))$, the later computed using image interpolation or approximation methods described in the next section. Thus methods for estimating the probabilities density functions $P_T(t)$, $P_{S_f}(s)$ and $P_{S_f,T}(s,t)$ are needed. Two families of methods are used for that purpose. The simplest relies on estimating

$P_{S_f,T}(s,t)$ using joint histograms of their intensity values [70]. To that end the intensity values of the images are divided into discrete, contiguous, and equally distributed intervals indexed using integers $i$ and $j$. Thus:

$$P_{S_f,T}(s,t) \cong P_{S_f,T}(i,j) = \frac{\text{number of co-ocurrences } i,j}{\text{total number of samples}}. \tag{2.32}$$

The marginal density functions are:

$$P_{S_f}(i) = \sum_j P_{S_f,T}(i,j), \tag{2.33}$$

and

$$P_T(j) = \sum_i P_{S_f,T}(i,j). \tag{2.34}$$

The integrals in equations (4.21),(2.29),(2.30), and (2.31) are then replaced by discrete sums over the bins of the histograms.

Joint histograms methods are the most computationally efficient, and most widely used, methods for estimating the mutual information similarity measure between two images. The method, however, is not void of shortcomings since, due to the discrete nature of histograms, small changes in transformation $f(\mathbf{x})$ can cause changes in the histograms that are difficult to predict, making estimates of the gradient of the mutual information with respect to the transformation parameters not well defined.

Methods for estimating the necessary density functions from the sample data using kernel functions [101] present an alternative to histogram methods. The idea is to build an approximation to the joint *pdf* as a linear combination of basis functions $G(s,t)$ centered at the data points $s_i, t_i$ (taken from images $S_f(\mathbf{x})$ and $T(\mathbf{x})$):

$$P_{S_f,T}(s,t) = \sum_{i=1}^{N} G(s-s_i, t-t_i). \tag{2.35}$$

The marginal *pdf*'s can then be estimated via integrating over separate dimensions, as above. The basis functions $G(s,t)$ are usually chosen to be radially symmetric, two-dimensional Gaussian functions:

$$G(s,t) = \frac{1}{\sigma_1 \sigma_2 2\pi} \exp\left(-\frac{1}{2}\left(\frac{s}{\sigma_1}\right)^2 - \frac{1}{2}\left(\frac{t}{\sigma_2}\right)^2\right) \tag{2.36}$$

though basis functions of short compact support have also been used [117].

While kernel based approaches overcome some of the difficulties associated with histogram methods (mainly differentiation of the cost function w.r.t. the spatial transformation) their computational complexity can be high. Thus researchers are forced to use a low number of data samples to estimate equation (2.35) and to use stochastic optimization approaches [56].

Finally, one novel and interesting class of image registration problems is the multi-channel (multi-spectral one). In this class of problems, one is interested in registering two or more images that at each spatial coordinate $\mathbf{x} \in \mathbb{R}^d$ contain not a scalar number by a vector of values each representing unique information. Multi-channel similarity measures have been proposed in Rohde *et al.* [89, 88]. Naturally, the computational load of multi-channel registration problems is significantly higher. However, it is hoped that the additional information present in the multiple channels can increase the accuracy of the spatial transformations obtained.

### 2.2.4 Image interpolation and approximation

Before any of the image similarity measures mentioned above can be computed, a continuous model for the discretely sampled image data is needed. That is, from a set of discrete samples $S(i,j,k) \in \mathbb{R}$, $1 \leqslant i \leqslant N, 1 \leqslant j \leqslant N, 1 \leqslant k \leqslant N$,

organized on a cartesian grid representing a 3D image (note that the number of samples in each dimension may differ), a mathematical operation is necessary for computing the image values at an arbitrary spatial position $\mathbf{x} = \{x, y, z\}^T \in \Omega$.

Because of practical computational requirements the continuous image is usually computed as a linear combination of separable basis functions:

$$S(\mathbf{x}) = \sum_{\mathbf{i} \in \mathbb{Z}^d i} c(\mathbf{i}) b(\mathbf{x} - \mathbf{i}), \tag{2.37}$$

where

$$b(\mathbf{x}) = \prod_{i=1}^{d} \hat{b}(x_i), \tag{2.38}$$

and $c(\mathbf{i}) \in \mathbb{R}$ are the coefficients of the expansion. From sampling theory it is known that a continuous version of a bandlimited signal (a continuous function whose Fourier transform has finite compact support) can be perfectly reconstructed from a sample version of the same signal if the sampling operation satisfied the Nyquist criterion. In this case $c(\mathbf{i}) = S(\mathbf{i})$ and

$$\hat{b}(x) = \mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}. \tag{2.39}$$

Sinc interpolation kernels, however, are seldom used in medical image analysis because of several reasons. First, it is often argued that the objects one is interested in reconstructing cannot be assumed to be bandlimited [107]. In addition, due to the infinite support of (2.39) equation (2.37) cannot be implemented, with the exception of when (2.37) is being used to compute pure translations of periodic signals. Lastly, due to its slow decay, image interpolation using the sinc function is know to produce so called 'ringing' artifacts.

To circumvent these and other difficulties, researchers often prefer to use apodized and truncated versions of (2.39) in computing spatially transformed

images $S(f(\mathbf{x}))$. In addition, simpler basis functions of compact support are also used. Amongst these, the linear (hat function) basis functions is by far the most widely used interpolation procedure used by image registration methods. The linear basis function takes the form:

$$\hat{b}(x) = \begin{cases} 1 - |x| & \text{if } |x| \leq 1; \\ 0 & \text{if } |x| > 1. \end{cases} \tag{2.40}$$

This is the most used interpolation kernel since its support is short, allowing for efficient computer implementations, and the coefficients in the expansion in (2.37) are the data values themselves: $c(\mathbf{i}) = S(\mathbf{i})$. Citing continuity and approximation properties, some researchers prefer higher order basis functions such as B-splines:

$$\hat{b}(x) = \beta^n(x) = \sum_{k=0}^{n+1} \frac{(-1)^k (n+1)}{(n+1-k)!k!} \left( \frac{(n+1)}{2} + x - k \right)_+^n, \tag{2.41}$$

where $x \in \mathbb{R}$ and $n \in \mathbb{N}$ represents the order of the spline. By definition $(x)_+^n = (\max(0, x))^n$. The coefficients of the expansion (2.37) can be computed by solving a linear system similar to the one set up in equation (2.2). Alternatively, the coefficients of the expansion may also be computed using recursive filtering techniques [114, 115].

Objective comparisons between the several image interpolation methods in the context of medical imaging (linear, truncated sinc, B-splines, etc.) have been performed in [66, 107, 108, 75]. Lehmann *et al.* perform quantitative and qualitative comparisons between several interpolation methods including nearest neighbor, linear, quadratic, piece-wise cubic polynomial, cubic B-spline, truncated sinc and others. Qualitative analysis was done by assessing interpolation error for particular interpolation tasks on digital X-ray image data. Quantitative analysis was

performed by comparing each function's Fourier properties and computational complexity in computer implementations. The authors conclude that piece-wise cubic polynomials can be recommended for most medical imaging operations for its computational efficiency, simple implementation, and continuity properties.

Thévenaz *et al* [107, 108] advocate the use of B-splines as a good compromise in terms of their linear approximation properties and computational complexity. Conclusions are based on experiments in which a single image is successively rotated by a certain amount of degrees until an approximation of the original image is obtained. The error between the original image and the successively rotated image is measured. The authors conclude that the approximation order of the basis function being used to generate the continuous image model is the most important index representing the quality of the interpolated images, while the support of the basis function is the most indicative parameter in terms of the computational requirements of using a particular basis functions. Meijering *et al.* [75] arrive at similar conclusions using similar experiments, though in this report effort is made to base results obtained using a variety of real medical image data.

An interesting effect of interpolation methods in the context of medical image registration is reported in [70, 85, 112, 117]. In these works authors report that the Mutual Information similarity measure $I(T(\mathbf{x}), S(f(\mathbf{x}))$, given in equation (4.21), can behave undesirably as the parameters of the spatial transformation $f(\mathbf{x})$. Such artifacts cause spurious local optima in the objective function, often restricting optimization methods (see next section) from finding the global optima and producing precise and accurate results. Though such artifacts have been detected for almost ten years now, and researchers have linked these artifacts to the interpolation or approximation procedure being used, their investigation of

the phenomenon is empirical at best and fail to describe the true nature of the artifacts and proper ways to mitigate them [70, 85, 112, 117].

### 2.2.5 Numerical optimization strategies

The global energy minimization approach to image registration described earlier is combined with a numerical optimization strategy to find the spatial transformation that aligns the images. Here again we differentiate parametric and non-parametric methods for deformation fields since the implementation of numerical optimization strategies can differ in practice. In both cases, the deformation field being sought belongs to the class of diffeomorphisms. In non-parametric deformation-based registration, however, optimization usually involves finite difference schemes to solve the derived Euler equation (2.13). In parametric deformation-based image registration, numerical methods usually involve traditional nonlinear optimization methods such as steepest descent, and Newton-type methods.

We first derive one very popular finite difference implementation, known as the optical flow formulation for the nonrigid matching problem, based on the variational setup described earlier. Let the functional being minimized be the squared of the 2-norm of the difference between the images:

$$\Theta(\mathbf{u}) = \int_{\Omega} F\left(S(\mathbf{x}+\mathbf{u}), T(\mathbf{x})\right) d\mathbf{x} \tag{2.42}$$

with $\mathbf{u}(\mathbf{x}) + \mathbf{x} = f(\mathbf{x})$ and

$$F\left(S(\mathbf{x}+\mathbf{u}), T(\mathbf{x})\right) = \left(S(\mathbf{x}+\mathbf{u}) - T(\mathbf{x})\right)^2 \tag{2.43}$$

A constraint term will be added later, but for the moment let us just consider the functional above. Following the discussion provided earlier, a stationary point in

(2.42) must satisfy the Euler-Lagrange equation (2.13). Setting the first variation of (2.42) to zero we have:

$$2\left(S(\mathbf{x}+\mathbf{u})-T(\mathbf{x})\right)\nabla S(\mathbf{x}+\mathbf{u})=0. \tag{2.44}$$

Solving (2.44) *per se* in the context of image registration is not necessarily a very good idea. For example, expanding constant value regions so that the entire domain $\Omega$ is dominated by regions in which $\nabla S(\mathbf{x}+\mathbf{u})\sim 0$ would go a long way towards satisfying (2.44) without necessarily aligning the images. Some of this difficulty can be overcome through constraints applied on the spatial transformation, soon to be discussed. Another option is to set

$$S(\mathbf{x}+\mathbf{u})-T(\mathbf{x})=0. \tag{2.45}$$

Using a first order expansion of the source image $S$ about coordinate $\mathbf{x}+\mathbf{u}$

$$S(\mathbf{x}+\mathbf{u})\sim S(\mathbf{x})+\mathbf{u}\cdot\nabla S(\mathbf{x}) \tag{2.46}$$

we have

$$\mathbf{u}\cdot\nabla S(\mathbf{x})=T(\mathbf{x})-S(\mathbf{x}). \tag{2.47}$$

Naturally, the equation above does not uniquely define $\mathbf{u}$. However, one may check by simple substitution that

$$\mathbf{u}=\frac{\left(T(\mathbf{x})-S(\mathbf{x})\right)\nabla S(\mathbf{x})}{\|\nabla S(\mathbf{x})\|^2} \tag{2.48}$$

does satisfy (2.47). The equation above addresses the concern expressed earlier with equation (2.44) since the magnitude of the gradient of $S(\mathbf{x})$ is eliminated. Naturally, the equation (2.48) is only valid for small displacements, since a first order Taylor expansion was used in deriving it. Therefore registration algorithms

use it iteratively by setting $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{u}_u$. It is worth noting that equation (2.48) is unstable for small values of $\|\nabla S(\mathbf{x})\|^2$ a difficulty which can be circumvented by using

$$\mathbf{u} = \frac{(T(\mathbf{x}) - S(\mathbf{x}))\,\nabla S(\mathbf{x})}{\|\nabla S(\mathbf{x})\|^2 + (T(\mathbf{x}) - S(\mathbf{x}))^2}. \tag{2.49}$$

In optical flow parlance $\mathbf{u}$ is considered as a velocity measurement because the images $S(\mathbf{x})$ and $T(\mathbf{x})$ are usually taken to be two successive frames in a time series of images.

Finally one must also constrain the method above since due to noise and flat regions in the images the iterative method above may become highly unstable. To that end the Lagrangian in (2.43) is substituted with:

$$F\left(S(\mathbf{x}+\mathbf{u}), T(\mathbf{x})\right) = \left(S(\mathbf{x}+\mathbf{u}) - T(\mathbf{x})\right)^2 + \|\mathbf{u}_\mathbf{x}\|^2 \tag{2.50}$$

where $\mathbf{u}_\mathbf{x}(\mathbf{x}) = \nabla \mathbf{u}(\mathbf{x})$ (note that in two or more dimensions, the gradient norms are component-wise). The new Euler equation becomes:

$$2\left(S(\mathbf{x}+\mathbf{u}) - T(\mathbf{x})\right)\nabla S(\mathbf{x}+\mathbf{u}) - 2\mathbf{u}_{\mathbf{xx}} = 0. \tag{2.51}$$

Considering the second term by itself, we must have that the component-wise Laplacian $\mathbf{u}_{\mathbf{xx}}$ be zero. It is possible to estimate the component-wise Laplacian $\mathbf{u}_{\mathbf{xx}}$ using a convolution with a Gaussian filter $G(\mathbf{x})$:

$$\mathbf{u}_{\mathbf{xx}} = G * \mathbf{u} - \mathbf{u}. \tag{2.52}$$

Thus, the equation (2.49) is usually regularized via convolution with a Gaussian kernel:

$$\mathbf{u} = \left(\frac{(T(\mathbf{x}) - S(\mathbf{x}))\,\nabla S(\mathbf{x})}{\|\nabla S(\mathbf{x})\|^2 + (T(\mathbf{x}) - S(\mathbf{x}))^2}\right) * G. \tag{2.53}$$

The derivation above was done to show that popular image registration methods such as the one presented in [109] can be viewed as finite difference solutions to the Euler equations associated with a global variational energy between the source and target images - in this case the 2-norm of their difference. Another popular derivation is based on the optical flow assumption that the intensity value of a particular structure in the image does not change, only its position [54]. The derived equations, however, are nearly identical. Variational approaches for minimization of similarity measures such as the correlation coefficient, and the mutual information have also been developed [50] and shown to be well-posed [35].

When the spatial transformation $f(\mathbf{x})$ is explicitly parameterized, say as a rigid-body transformation as in equation 2.14, or as a combination of basis functions as in 2.15, the optimization is performed with respect to the model parameters themselves. For example, in the case of B-spline based transformation models such as the one defined through equations (2.15) and (4.34), to goal is to seek the value of a set of coefficients $\mathbf{C}$ of $N$ vector coefficients $\mathbf{c}_i$ such that the overall cost function $\Theta$ expressed in (2.7) (similarity measure plus constraint term) is minimized. The set of basis function coefficients $\mathbf{C}$ is usually initialized to zero (unless prior information is available) and it is updated using the following procedure:

$$\mathbf{C}_{k+1} = \mathbf{C}_k - \alpha_k \xi \left( \mathbf{C}_k \right), \tag{2.54}$$

where $\xi(\cdot)$ is a function that determines a descent direction, and the constant $\alpha_k$ is determined through a line minimization procedure. In the simplest example, and often the most used in biomedical image registration, the descent direction is chosen as the steepest gradient, $\xi(\mathbf{C}) = \nabla\Theta(\mathbf{C}_k)$, however many other options such as the conjugate gradient method, quasi-newton methods, etc., are available

[77]. The procedure is iterated until

$$\|\nabla\Theta(\mathbf{C}_k)\| \leqslant \epsilon, \tag{2.55}$$

or

$$|\Theta(\mathbf{C}_k) - \Theta(\mathbf{C}_{k-1})| \leqslant \epsilon, \tag{2.56}$$

for arbitrary $\epsilon$.

Naturally, the basis function-based approach described above is very computationally expensive, especially for performing inter-patient nonrigid registration in three dimensions. Many structures of interest in medical images, especially in the brain, are in the order of millimeters. Deforming such structures requires placing basis functions at approximately every couple of millimeters which can require the optimization of a few hundred thousand basis function coefficients. For example, registering two typical three-dimensional MR image volumes (256x256x128 pixels large) using a 64x64x32 regular grid of splines generates a 393 216-dimensional search space. Finding an optimum in such a search space is not only time consuming but difficult because of convergence to local optima.

By necessity, methods for accelerating nonrigid registration methods based on parametric deformation models have been extensively researched. The first approach is based on finding the optimal transformation through successive approximations computed within a multi-resolution image model. That is, a solution $f_k(\mathbf{x})$ the problem defined by equations (2.5) and (2.7) is computed using a coarse (low resolution) approximation of the images $S(\mathbf{x})$ and $T(\mathbf{x})$. A coarse approximation of the images allows for faster computation of the overall cost function, as well as its gradient with respect to the transformation parameters. Thus, an initial estimate $f_k(\mathbf{x})$ can be computed relatively quickly. This estimate

is then used to restart the optimization using a higher resolution approximation of the images and compute a new estimate of the spatial transformation $f_{k+1}(\mathbf{x})$.

The coarse to fine optimization method described above relies on building multi-resolution pyramid-type approximations to the image data. To date, many methods for building such pyramids are available. One of the most used, due to its simplicity and self-consistency, is the multi-resolution approximation scheme explained in [116]. The idea is to use a hierarchical approximation over nested B-spline subspaces. Let $\mathrm{V}^Q \subset \mathrm{V}^{Q-1} \subset \cdots \subset \mathrm{V}^0$ define a sequence of $Q+1$ nested subspaces, with each $\mathrm{V}^q$ defined as the span of vectors:

$$\mathrm{V}^q = \mathrm{Span}\{\beta^n \left(\frac{\mathbf{x}}{2^q} - \mathbf{i}\right)\}, \tag{2.57}$$

where $\beta^n(x)$ is the $n^{\text{th}}$ order b-spline function defined in equation (2.41). The coefficients of the expansion at each level are computed by minimizing the 2-norm between the original signal and its approximation. This computation can be performed efficiently via recursive filtering operations [114, 115].

The multi-resolution image approximation scheme discussed above can also be combined with multi-scale approximations to the spatial transformations $f(\mathbf{x})$, when $f(\mathbf{x})$ is also being modeled as a combination of basis functions. That is, when using low resolution approximations of the images $S(\mathbf{x})$ and $T(\mathbf{x})$ it does not make much sense in using dense models for $f(\mathbf{x})$ since, due to the coarse representation of the image data, highly localized deformations are not likely to be picked up. Thus, low resolution approximations of the image data are coupled with low resolutions approximations of the spatial transformation. The solutions obtained using low resolution image and spatial transformation models are then used to restart the problem using higher resolution data and

transformation approximations.

Finally, additional speed up may be achieved through selective placement and optimization of the basis functions that define the spatial transformation model. Using the notion that not all regions in the domain $\Omega$ may benefit from non-rigid displacements (i.e. in some regions the images may be severely mismatched, while in others less so), Rohde *et al.* [87] use a simple method for detecting which regions will benefit most from spatial adjustment, and focus most computations in these regions. The optimization is done within the multi-resolution, multi-scale context described above. Results show that such adaptive deformation field modeling approaches can significantly speed up computations without compromising the quality of results.

# Chapter 3

# Overview of Contributions

As shown in the previous chapter, the vast majority of automatic, intensity-based, image registration methods can be characterized as specific examples of global energy minimization strategies. That is, the goal in image registration problem is to find the spatial transformation $f(\mathbf{x})$ that minimizes a cost function containing an image dissimilarity measure term and a constraint term. Registration methods based on explicit parametric models of the spatial transformation $f(\mathbf{x})$ use traditional optimization approaches, such as the steepest gradient descent method and others, to find the transformation parameters that optimize the cost function of interest. It was also shown that image registration methods based on partial differential equations formulations, such as the optical flow method, can also be viewed within an energy minimization context. This is because the equations used in such methods can be shown to be related to the Euler-Lagrange equations that arise from a global variational energy framework.

In this chapter the original contributions to the field of registration methods for quantitative imaging described in this dissertation are reviewed in the context of the energy minimization framework presented in the previous chapter. It is shown that much can be gained by including a model for the image formation

process in accounting for artifacts caused by patient (object) motion, device dependent geometric distortions, and system noise in the design and implementation of post-acquisition image processing methods.

We start by using a general shift invariant image formation equation, based on which we calculate the effects of patient motion and geometric distortions. Due to uncertainties introduced by thermal noise in several stages of modern imaging systems, the model must be statistical in nature. Next, we highlight the use of the image formation model for addressing some of the previously 'open' problems in biomedical image registration. The individual aspects of the research appear in subsequent publications, included here as separate chapters.

## 3.1 Image formation model

Until now the images of interest have been described as a set of discrete values sampled during an imaging experiment and stored digitally in computer memory. The sampled images are then registered using an energy minimization framework together with a continuous model for the images. In many applications, however, it is useful to note that the digital images being manipulated in the computer are not necessarily an exact representation of the functions one wishes to obtain (i.e. a function describing the properties of the object of interest with infinite fidelity) but, rather, they represent the desired function as viewed through a particular imaging instrument and experiment. We consider here shift-invariant imaging systems of the form:

$$S(\mathbf{x}) = \int W(\mathbf{p}) h(\mathbf{x} - \mathbf{p}) d\mathbf{p} \tag{3.1}$$

where $\mathbf{x}, \mathbf{p} \in \mathbb{R}^d$, with $d$ being the dimensionality of the images, $W(\mathbf{x})$ representing the object or scene from which one wishes to obtain information, and $h(\mathbf{x})$ is the transfer function of the imaging system. While this model for image formation may not be universal, it is a good approximation for many imaging devices. Optical based systems such as charged coupled devices, for example, are often modeled using a linear, shift invariant, image formation equation as the one above. As shown in chapter 5, in the appendix, magnetic resonance imaging systems can also be viewed in this framework. Even though not discussed in this dissertation, the model above is also suitable to describe, at least approximatelly, other imaging systems such as synthetic aperture radar and x-ray computed tomography [78].

Because of patient motion with respect to the imaging instrument, we must realize that a particular instance of the image $S(\mathbf{x})$ may be corrupted by a function $f(\mathbf{x})$ describing the motion of the object:

$$S(\mathbf{x}) = \int W(f(\mathbf{p}))h(\mathbf{x} - \mathbf{p})d\mathbf{p}. \tag{3.2}$$

Note that the integration process above is assumed to be infinitesimally short (in time) compared to any significant time variations in the motion function $f$. Naturally, this is not always a reasonable assumption, since significant motion can occur during lengthy image acquisitions. Motion during data acquisition, however, causes defocusing of the image in a way that cannot be corrected using image domain-based registration methods.

In addition to patient motion, the acquired image $S(\mathbf{x})$ may also suffer from geometric distortions represented here by $g(\mathbf{x})$ so that:

$$S(\mathbf{x}) = \int W(f(\mathbf{p}))h(\mathbf{x} - g(\mathbf{p}))d\mathbf{p}. \tag{3.3}$$

It is important to note that any image formation process is statistical in nature due to the fact that noise from thermal and other effects is almost always present. Thus the measurement at location $\mathbf{x}$ is corrupted by a zero mean random variable $\eta(\mathbf{x})$ so that:

$$S(\mathbf{x}) = \int W(f(\mathbf{p}))h(\mathbf{x} - g(\mathbf{p}))d\mathbf{p} + \eta(\mathbf{x}). \tag{3.4}$$

Finally, we note that the function above is usually sampled at discrete positions $\mathbf{x}_{i,j,k}$ with $i, j, k \in \mathbb{N}$ usually defining a regular grid. Digital sampling, strictly speaking, is a nonlinear operation since most analog to digital converters effectively implement step-like functions. Thus, any digital image formation process is inherently nonlinear. In many cases, however, the effects of digital sampling can be minimized in real experiments, the effects of sampling error can often be taken into account stochastically. That is, the error due to digital sampling in the measured signal $S(\mathbf{x})$ can be viewed as probabilistic and modeled in the term $\eta(\mathbf{x})$. In some cases, the image above may also be convolved with an additional filter during reconstruction (such as in 'filtered' back projection reconstruction in computed tomography, or appodisation filters in magnetic resonance images). In this case, the random variable above is an example of a correlated stochastic process.

Now, let $\mathbf{y} = \mathbf{x} - g(\mathbf{p})$. The equation above can be rewritten as:

$$S(\mathbf{x}) = \int W(f(g^{-1}(\mathbf{x} - \mathbf{y})))h(\mathbf{y})d\mathbf{p} + \eta(\mathbf{x}). \tag{3.5}$$

with $g^{-1}(\mathbf{x})$ representing the function inverse of the spatial transformation $g(\mathbf{x})$. Naturally, for the equation above to make sense, $g(\mathbf{x})$ must be a homomorphism. Let $|J(g^{-1}(\mathbf{u}))|$ represent the determinant of the Jacobian matrix of the spatial transformation $g^{-1}(\mathbf{u})$. In this case we can integrate with respect to $\mathbf{y}$ instead of

**p**:

$$S(\mathbf{x}) = \int W(f(g^{-1}(\mathbf{x}-\mathbf{y})))h(\mathbf{y})|J(g^{-1}(\mathbf{x}-\mathbf{y}))|d\mathbf{y} + \eta(\mathbf{x}). \tag{3.6}$$

If we think of the transfer function of the imaging system $h(\mathbf{y})$ as an approximation to the Dirac delta distribution, the measured image can be approximately expressed as:

$$S(\mathbf{x}) = \tilde{W}(f(g^{-1}(\mathbf{x})))|J(g^{-1}(\mathbf{x}))| + \eta(\mathbf{x}). \tag{3.7}$$

$\tilde{W}$ above is meant to represent the object function blurred by the point spread function ($psf$) of the imaging system. The equation above states that, for the case of shift-invariant imaging systems under the effects of motion and geometric distortions, the measured image can be thought of as, approximately, the object shifted by the function $f(g^{-1}(\mathbf{x}))$, weighted by the determinant of the Jacobian matrix of the transformation $g^{-1}$, convolved with the point spread function of the imaging system, and with noise added. The function $|J(g^{-1}(\mathbf{x}))|$ represents the signal modulation caused by changes in volume (area) of the perceived object due to spatial distortion fields. Note that object motion (even non-rigid) does not have such an effect.

To remove the effects of motion and distortion in the measured image $S(\mathbf{x})$ one must find $\tilde{f}^{-1}$ (the inverse of the displacement field due to motion) and $\tilde{g}$ so that:

$$S(g(f^{-1}(\mathbf{x})))|J(g(\mathbf{x}))| \sim \tilde{W}(\mathbf{x}) + \tilde{\eta}(\mathbf{x}). \tag{3.8}$$

Note that the noise variable above $\tilde{\eta}(\mathbf{x})$ differs from the noise variable in the 'un-processed' image defined in equation (3.7). This is because the spatial transformation defined above needs to be computed based on a finite set of samples from the original image, using an approximation or interpolation scheme. Let $\mathbf{S}$ represent the vector of image values sampled from (3.7) using a regular grid-type

sampling scheme. Let $\Sigma_{\mathbf{S}}$ represent the covariance matrix of the measurements. As shown in the previous chapter, image interpolation or approximation schemes can usually be computed as a linear operation on the sampled image data values. Let **AS** define such an operation. The covariance matrix of the transformed image is then given by $\mathbf{A}\Sigma_{\mathbf{S}}\mathbf{A}^T$. More precise explanations of the effects of the processing steps necessary to correct for the artifacts above on several steps in a quantitative imaging pipeline will be described in the following chapters.

The remaining chapters of this dissertation consist of separate journal publications. In each of these, the stochastic, linear, image formation model described above, combined with different modeling and optimization methods, is used to tackle different aspects of image registration problems. A brief description of the content of each publication is described here.

## 3.2 Measuring image similarity

Here the effects of thermal noise in the image formation process on the functional minimization approach to image registration are examined. It is shown that certain types of polynomial interpolation methods, in particular low order ones, used in computing spatially transformed versions of the images will cause undesired effects on the image dissimilarity measures typically used in image registration. Previously, researchers have reported that image interpolation methods can cause local optima in mutual information-type image similarity measures [70, 85, 112, 117]. These articles, however, merely present an empirical study of the effects of different interpolation methods on the mutual information similarity measure with respect to global image translation. We show that such artifacts are general, in the sense that they are not limited to mutual information

similarity measures but will also occur in $L_2$ and correlation-type cost functions. Moreover, they are also not limited to global transformations, such as translation and rotations, but also occur in more local, nonlinear type of transformations. We explain the source of the artifacts in terms of the covariance properties of spatially transformed images. We explain why some interpolators such as the linear hat function and low order B-splines will cause local optima artifacts in popular similarity measures and argue that sinc approximating basis functions should be used instead. In some simple analytical examples we give analytical formulas describing the locations and shape of the local optima. We also validate our hypotheses using real and simulated image data. Our results contradict several previous studies which found no practical evidence for using higher-order sinc approximation functions [66, 107, 108, 75].

## 3.3 Registration of diffusion weighted MRI's

Here the image formation model defined in equation (3.7) is used to derive a comprehensive approach for correcting artifacts related to patient motion and eddy current-induced geometric distortions in diffusion weighted MRI experiments of the human brain. MRI-based studies of microscopic water molecule displacement [17] have become a valuable tool in the quantification of many important physical properties of biological tissues, *in vivo*. They can indicate the direction of nerve bundles [83], for example, amongst many other important quantities. Previous approaches for addressing the problem of geometric distortions caused by the fast switched diffusion weighting magnetic field gradients [47, 55, 73] exist. However, most of the approaches are *ad-hoc* in the sense that little or no justification is given for the models being used. Moreover, the extent of all sources of artifacts,

including motion, distortion, intensity modulations, etc., are not considered. We model and optimize patient motion and geometric distortions simultaneously, as suggested by the model in equation (3.7) in three dimensions. A 3D rigid body transformation is used to account for patient motion, while a spherical harmonics expansion in cartesian coordinates is used to model the eddy current-induced fields that cause geometric distortions. In addition, intensity and b-matrix corrections (see chapter 5) are performed. The algorithm is implemented numerically using a mutual information-type cost function and a gradient descent-type minimization strategy. The series of images is produced using a single interpolation step, with minimum degradation of image quality. Results with real data show that the approach can recover much of the diffusion information lost due to spatial misalignment in typical diffusion tensor imaging experiments. Note that most of the work presented in this chapter was previously published (with modifications) in Rohde *et al.* [91].

## 3.4 Motion correction in optical mapping

Here similar techniques are used to address motion artifacts in optical mapping experiments. High-resolution optical mapping is an emerging technique to record the activation and propagation of transmembrane potential on the surface of cardiac tissues. Important electrodynamic information previously not available from extracellular electric recordings can be extracted from these detailed optical recordings. A typical imaging setup is shown in figure 3.1. Fluorescence from the potentiometric dye-stained heart surfaces is elicited by a solid state laser and the induced fluorescence is imaged with a high-speed CCD camera. The data from the experiment is then used to analyze the spatiotemporal distribution of

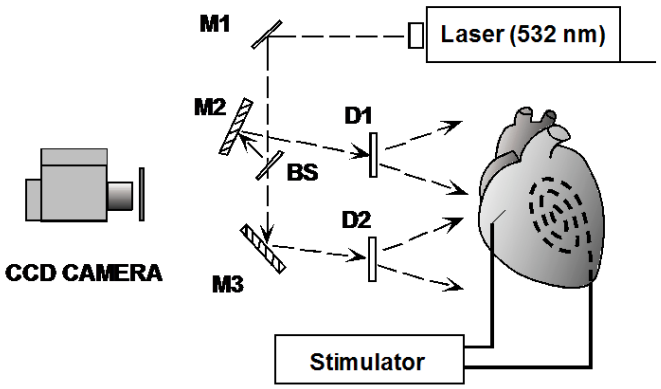


Figure 3.1: Typical cardiac optical mapping experiment setup.

transmembrane potential under different stimuli patterns. However, motion from the beating heart remains one of the main sources of artifacts in such experiments. It can can significantly degrade the quality of electro-physiological measurements such as action potential duration. A novel approach for minimizing such artifacts based on image registration methods is presented. The image registration method is an alternative to more traditional approaches such as mechanical restraint of the heart of addition of chemical uncouplers, which can interfere with the phenomena being measured. Finally, note that the work presented in this chapter was published, with modifications, in Rohde *et al.* [93].

## 3.5 Post-registration noise variance estimates

As fitting and estimation procedures from registered image data become increasingly more elaborate and quantitative, knowledge of the intensity variance due to noise will become more important for increasing the accuracy and scientific value of the results obtained from them. A Method for estimating the variance

in registered images is presented. The general approach is a simple extension of the stochastic image model described above and can be summarized as follows. The output of the registration procedure is computed using an image interpolation or approximation procedure. More often than not, the interpolation or approximation procedure can be written as a linear combination of the values of the images being registered. The coefficients of the linear combination are determined by the choice of interpolation or approximation kernel. Since the values of the images being registered are typically corrupted by noise, this operation can be viewed as a linear combination of random variables. The variance of the linear combination is given by well-known statistical formulas. It is shown that incorrect information about intensity variance due to noise can significantly degrade the quality of quantitative results resulted from imaging experiments. Particular focus is given to analyzing diffusion tensor imaging experiments. The results, however, should be general and apply to almost any situation in which registered or processed images are used for quantitative experiments. Finally, though modifications and improvement were made, most of the work presented in this chapter was previously published in Rohde *et al.* [92].

# Chapter 4

# Measuring image similarity for image registration

## 4.1 Introduction

Image registration is the process of identifying the spatial correspondence between different images. Registration of medical images is an important procedure in many aspects of biomedical research and clinical practice where it is used to fuse information from images of a single subject taken at different times to account for subject motion, or geometric distortions. Image registration methods are also used by neuroscientists to relate images taken from different subjects. In the context of brain neuroanatomy for example, researchers (neuroscientists assisted by imaging engineers, scientific computation experts, mathematicians, and statisticians) have long used image registration methods to study the variation of biological tissue properties, such as shape and composition, described in images across a given population. The aim in such endeavors is usually to provide a quantitative description of a healthy 'normal' population, as opposed to a diseased one.

Though implementations vary, most methods seek to solve the digital image registration problem within an optimization framework whose goal is to find a

function $f\colon \mathbf{x} \rightarrow \mathbf{x}'$ that transforms the spatial coordinates $\mathbf{x} \in \mathbb{R}^d$ of a target image $T(\mathbf{x})$ to the spatial coordinates $\mathbf{x}' \in \mathbb{R}^d$ of a source image $S(\mathbf{x}')$. For the source and target images to be spatially aligned the mapping function $f$ should be chosen in such a way as to optimize some cost function (objective function) between the two images. Mathematically, the image registration problem can be stated as a minimization problem:

$$\min_f \Theta(S(f(\mathbf{x})), T(\mathbf{x}), f) \tag{4.1}$$

where $\Theta(\cdot, \cdot, \cdot)$ represents the objective function being optimized. To prevent registration methods from producing transformations that violate the intrinsic topology of the image the objective or other basic constraints function $\Theta$ is usually computed as a sum of two terms:

$$\Theta(S(f(\mathbf{x})), T(\mathbf{x}), f) = I(S(f(\mathbf{x})), T(\mathbf{x})) + C(f). \tag{4.2}$$

The term $C(f)$ is a 'regularization' term meant to prevent wild spatial oscillations in the transformation $f$. Popular choices for regularization of registration methods include $C(f) = \int |Df(\mathbf{x})|^2 d\mathbf{x}$ and $\int |\det(J(f(\mathbf{x})))|^{-1} d\mathbf{x}$, where $\det(J(f))$ stands for the determinant of the Jacobian matrix of $f$ and $D$ stands for a differential operator chosen based on arguments from continuum mechanics.

The function $I$ represents some distance measure, or equivalently, the negative of a similarity measure between the images being registered. The choice of $I$ is usually determined by the requirements of the application. In instances when images $S$ and $T$ are expected to be nearly identical under optimal alignment, $L_2$-norms or the sum of squared differences are often employed. When the relationship between the intensity values of $S$ and $T$ is unknown but expected

to be linear $I$ can be based on the correlation coefficient of their intensity values. In more complicated situations many researchers have shown the Mutual Information [61] similarity measure to be a good choice for $I$.

Before an estimate for the similarity measure $I(\cdot, \cdot)$ can be computed for an arbitrary $f(\mathbf{x})$, however, a suitable strategy for computing $S(f(\mathbf{x}))$ from a set of discrete samples $S(\mathbf{i})$, $\mathbf{i} \in \mathbb{Z}^d$ is necessary. Almost unanimously researchers choose to make image $S(\mathbf{i})$ continuous by modeling it as a linear combination of symmetric basis functions [66, 107, 75] determined by interpolation or approximation from the discrete data $S(\mathbf{i})$. In recent years, several researchers have reported that popular methods used to estimate Mutual Information, $L_2$ and correlation based similarity measures can behave unexpectedly with respect to the spatial transformation $f$ being applied [10, 85, 112]. Such artifacts have been attributed to the interpolation or approximation strategies being used. In this work we show that such interpolation artifacts occur in $L_2$ and correlation-based similarity measures whenever significant noise is present. We show that such artifacts stem from the fact that in estimating the value of the similarity measure for an arbitrary spatial transformation $f$ one is obliged to interpolate or approximate noisy data. The causes of interpolation artifacts encountered in mutual information registration curves are more complex in that system noise is not the only factor. However, like in the $L_2$ and correlation cases, artifacts in mutual information registration curves can also be explained by loss of spatial frequency information in the interpolated images. Naturally, the frequency content of interpolated images is better preserved when using sinc-based interpolation. As results show, interpolation artifacts in image similarity measures can be significantly reduced, often completely eliminated, by using sinc-approximating basis

functions.

The remaining is organized as follows. Using the theory of random processes we first describe the effects of spatial transformations on the variance and covariance structure of the interpolated image. We explain that noise covariance distortions can have undesired effects on measures of image similarity, such as the $L_2$ norm and correlation coefficient, and suggest ways through which these effects can be minimized. We also show how similar strategies can be used to alleviate so called 'grid' artifacts in mutual information curves. Finally, we test our solutions using both simulated and real magnetic resonance imaging (MRI) data.

## 4.2 Theory

### 4.2.1 Covariance properties of interpolated signals

Borrowing the approach described in [12] we use the following linear, stochastic, image model in our analysis:

$$S(\mathbf{x}) = \int W(\mathbf{p})\Upsilon(\mathbf{x}-\mathbf{p})d\mathbf{p} + e(\mathbf{x}) \tag{4.3}$$

where $\mathbf{x}$ and $\mathbf{p} \in \mathbb{R}^d$, $W(\mathbf{p})$ corresponds to the function describing the object being imaged, and $\Upsilon(\mathbf{x})$ is the point spread function of the imaging system. Note that throughout this paper all quantities will be assumed to be real valued. Unless noted otherwise, all integrals shall be evaluated from $-\infty$ to $\infty$. $e(\mathbf{x})$ refers to a zero mean stochastic process whose covariance structure will soon be described. We shall consider the integral part of equation (4.3) to be deterministic. The covariance function of the random process (4.3) is:

$$
\begin{aligned}
R_S(\mathbf{x}_1, \mathbf{x}_2) &= \mathrm{Cov}\{S(\mathbf{x}_1), S(\mathbf{x}_2)\} \\
&= \mathrm{E}\{(S(\mathbf{x}_1) - \overline{S}(\mathbf{x}_1))(S(\mathbf{x}_2) - \overline{S}(\mathbf{x}_2))\} \\
&= \mathrm{E}\{e(\mathbf{x}_1)e(\mathbf{x}_2)\} \\
&= \mathrm{Cov}\{e(\mathbf{x}_1), e(\mathbf{x}_2)\} = R_e(\mathbf{x}_1, \mathbf{x}_2)
\end{aligned}
$$

where

$$
\mathrm{E}\{e(\mathbf{x})\} = \overline{e}(\mathbf{x}) = \int e(\mathbf{x})\mathrm{pr}[e(\mathbf{x})]\ de(\mathbf{x}), \tag{4.4}
$$

and $\mathrm{pr}[e(\mathbf{x})]$ stands for the probability density function of the quantity $e(\mathbf{x})$. Thus, without loss of generality, for the purposes of analyzing the covariance of interpolated signals we momentarily assume that $S(\mathbf{x})$ is a zero-mean random process (i.e. $S(\mathbf{x}) = e(\mathbf{x})$ through subtraction of the deterministic part of (4.3)).

As stated above, before a given similarity measure $I(S(f(\mathbf{x})), T(\mathbf{x}))$ can be evaluated for an arbitrary spatial transformation $f$, a continuous model for the digital image $S(\mathbf{i})$ is needed. Most often researchers choose to model $S(\mathbf{i})$ as a linear combination of symmetric basis functions $h(\mathbf{x})$:

$$
S^c(\mathbf{x}) = \sum_{\mathbf{i} \in \mathbb{Z}^d} S(\mathbf{i})h(\mathbf{x} - \mathbf{i}). \tag{4.5}
$$

In this case we have chosen the coefficients of the linear combination to be the sampled image values. Note that the summations above are carried from $-\infty$ to $+\infty$. In our analysis we assume that all image samples needed for computing (4.5) are available. This is often true since most practitioners tend to use basis functions with short support. Moreover, in most cases the computations of interest are confined to an inner subset of the field of view of image $S(\mathbf{x})$. Thus,

in many cases, image values outside the field of view over which the image was originally sampled are not needed. In situations where the support of $h(\mathbf{x})$ is sufficiently large, thus requiring image samples outside the field of view of the image, the image can be extended by zero padding or by using periodic or mirror boundary conditions. Note that when sampled signals are extended, either by zero padding or by using reflections of the data, their covariance structure $R_S(\mathbf{x}_1, \mathbf{x}_2)$ must be extended the same way.

If we would like $S^c(\mathbf{i}) = S(\mathbf{i})$, the basis function $h(\mathbf{x})$ must obey the following:

$$h(\mathbf{i}) = 0 \quad \forall\ \mathbf{i} \in \mathbb{Z}^d \neq 0 \text{ and}$$
$$h(0) = 1.$$

Note that when $d > 1$, the interpolation kernel is generally taken to be separable:

$$\hat{h}(\mathbf{x}) = \prod_{j=1}^{d} h(x_j). \tag{4.6}$$

As shown in [4], expression (4.5) can be interpreted as a continuous filtering operation of the initial sampled image values

$$S^c(\mathbf{x}) = \int S_\delta(\mathbf{q}) h(\mathbf{x} - \mathbf{q}) d\mathbf{q} = (h * S_\delta)(\mathbf{x}), \tag{4.7}$$

where the notation $a_\delta$ represents the distribution consisting of the train of weighted Dirac delta impulses

$$a_\delta(\mathbf{x}) = \sum_{\mathbf{i}=-\infty}^{+\infty} a(\mathbf{i}) \delta(\mathbf{x} - \mathbf{i}). \tag{4.8}$$

We now compute the covariance structure of the continuous function (4.5) $R_{S^c}(\mathbf{x} + \Delta\mathbf{x}, \mathbf{x}) = \mathrm{Cov}\{S^c(\mathbf{x} + \Delta\mathbf{x}), S^c(\mathbf{x})\}$:

$$
\begin{aligned}
R_{S^c}(\mathbf{x}+\Delta\mathbf{x},\mathbf{x}) &= \mathrm{E}\{S^c(\mathbf{x}+\Delta\mathbf{x})S^c(\mathbf{x})\} \\
&= \mathrm{E}\{\int S_\delta(\mathbf{q}_1)h(\mathbf{x}+\Delta\mathbf{x}-\mathbf{q}_1)d\mathbf{q}_1 \int S_\delta(\mathbf{q}_2)h(\mathbf{x}-\mathbf{q}_2)d\mathbf{q}_2\} \\
&= \int\int h(\mathbf{x}+\Delta\mathbf{x}-\mathbf{q}_1)R_{S_\delta}(\mathbf{q}_1,\mathbf{q}_2)h(\mathbf{x}-\mathbf{q}_2)d\mathbf{q}_1 d\mathbf{q}_2,
\end{aligned}
$$

where $R_{S_\delta}(\mathbf{q}_1,\mathbf{q}_2) = \mathrm{E}\{S_\delta(\mathbf{q}_1)S_\delta(\mathbf{q}_2)\}$. If we assume that the correlations in the random process $S_\delta(\mathbf{x})$ are of such short range that $R_{S_\delta}(\mathbf{q}_1,\mathbf{q}_2)$ can be approximated by $z_\delta(\mathbf{q}_1)\delta(\mathbf{q}_1-\mathbf{q}_2)$, where $z_\delta(\mathbf{x}) = \mathrm{Var}\{S_\delta(\mathbf{x})\}$, then

$$
\begin{aligned}
R_{S^c}(\mathbf{x}+\Delta\mathbf{x},\mathbf{x}) &= \int z_\delta(\mathbf{q}_1)h(\mathbf{x}+\Delta\mathbf{x}-\mathbf{q}_1)h(\mathbf{x}-\mathbf{q}_1)d\mathbf{q}_1 \\
&= \sum_{\mathbf{i}=-\infty}^{+\infty} z(\mathbf{i})h(\mathbf{x}+\Delta\mathbf{x}-\mathbf{i})h(\mathbf{x}-\mathbf{i}).
\end{aligned}
$$

Moreover, if we are dealing with a constant variance random process with variance $z_\delta(\mathbf{x}) = \sigma^2$, the variance of the interpolated signal $S^c(\mathbf{x})$ is given by:

$$
\mathrm{Var}\{S^c(\mathbf{x})\} = R_{S^c}(\mathbf{x},\mathbf{x}) = \sigma^2 \sum_{\mathbf{i}=-\infty}^{+\infty} [h(\mathbf{x}-\mathbf{i})]^2. \tag{4.9}
$$

The purpose of the exercise above is to characterize the effects of the signal interpolation model (4.5) on the covariance structure of the signal. To illustrate this effect we have plotted equation (4.9), with $\sigma^2 = 1$, for several interpolating basis functions currently used in medical imaging in figure 4.1. See appendix A for the definition of the interpolators used. The truncated sinc basis function was computed using window width $W = 6$, as described in appendix A. As shown in figure 4.1, given a constant variance 'white' discrete random process, the continuous model expressed in (4.5) produces a function whose variance at each point in space is no longer uniform. This is especially true for low order interpolators such

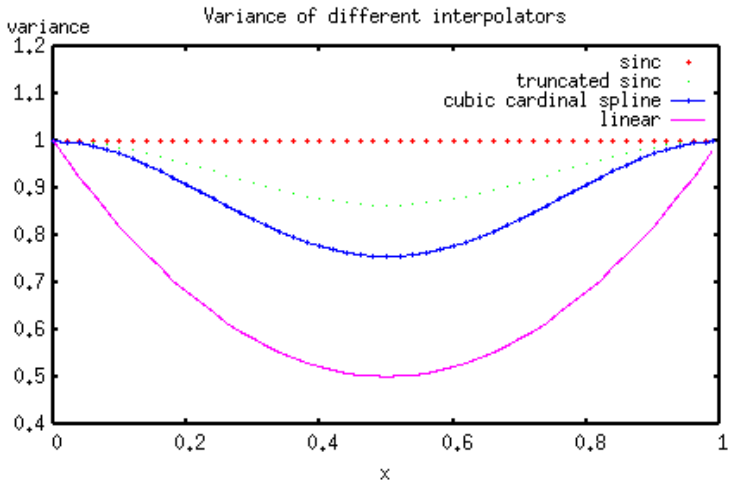


Figure 4.1: Plot of equation (4.9), with $\sigma^2 = 1$, for several interpolating basis functions typically used in medical imaging.

as the linear 'hat' function, but in general it is also true for any kind of interpolating basis function other than sinc. To understand why sinc interpolation preserves variance in this case it is enough to verify that $\sum_{\mathbf{i}=-\infty}^{+\infty}[\mathrm{sinc}(\mathbf{x}-\mathbf{i})]^2 = 1 \;\forall\; \mathbf{x}$ (see appendix B). Thus the covariance structure of a digital image that undergoes a geometric transformation via

$$S^c(f(\mathbf{x})) = \sum_{\mathbf{i}\in\mathbb{Z}^d} S(\mathbf{i})h(f(\mathbf{x}) - \mathbf{i}) \tag{4.10}$$

is dependent on the function $f$:

$$R_{S^c}(f(\mathbf{x}_1), f(\mathbf{x}_2)) = \int\int h(f(\mathbf{x}_1) - \mathbf{q}_1)R_{S_\delta}(\mathbf{q}_1, \mathbf{q}_2)h(f(\mathbf{x}_2) - \mathbf{q}_2)d\mathbf{q}_1 d\mathbf{q}_2. \tag{4.11}$$

If we are dealing with an approximately 'white' random process the variance of the interpolated signal is:

$$\mathrm{Var}\{S^c(f(\mathbf{x}))\} = \sigma^2 \sum_{\mathbf{i}=-\infty}^{+\infty} [h(f(\mathbf{x}) - \mathbf{i})]^2. \tag{4.12}$$

To illustrate this concept we have computed the following simulation. A series of 200 digital images was created using a random number generator such that the mean of each pixel was zero and the variance one, while the correlation between any two pixels was zero. The sample variance (in each pixel) is displayed on the left panel of figure (4.2). As expected, this image is fairly uniform depicting a constant-variance random process. Next, each of the 200 simulated images was rotated about its center by 4 degrees counter clock-wise using bilinear interpolation. The sample variance of the newly created series of rotated images was computed for each pixel and is displayed in the center panel of figure (4.2). As can be expected, the variance becomes non-uniform as a function of image coordinate and acquires a certain 'striped' configuration. Lastly, the original series of 200 images of random noise was again rotated about its center counter clock-wise by 4 degrees. This time, however, truncated sinc interpolation (with $W = 30$) was used. The sample variance of the rotated image series for each pixel is displayed on the right panel of figure (4.2). As can be expected, the variance of the rotated image series using truncated sinc interpolation is almost perfectly uniform.

### 4.2.2 Optimization of $L_2$-based similarity measures

We have shown above how the covariance properties of digital images are modified according to the spatial transformation $f$ being used during registration. We now show that this dependence can be detrimental to the image registration process for commonly used similarity measures such as the $L_2$ norm and the linear correlation coefficient. In order to facilitate the analysis we use the following vector notation for the source and target digital images. That is, $\mathbf{S} = \{S_0, S_1, \cdots, S_N\}^T$, with $N = m \times n$ where $m, n$ are the dimensions of the two-dimensional image. Using

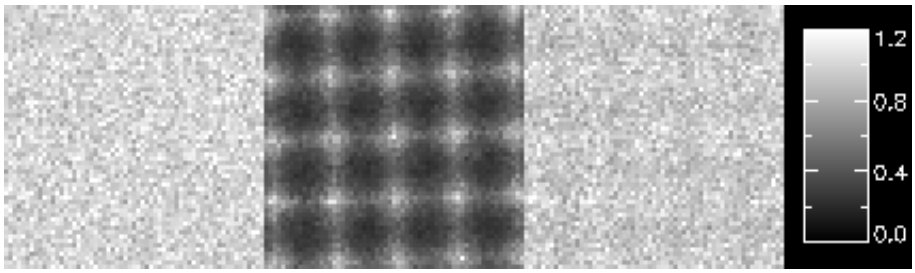


Figure 4.2: Variance of a stationary random process before and after rotation using different interpolators. The panel on the left shows the sample variance of the random process prior to any transformation. The middle panel show the variance of the random process after rotation of the images about their center using bilinear interpolation. The panel on the right shows the variance of the random process after the same rotation of the images this time computed using sinc approximating basis functions.

(4.3) we can write $\mathbf{S} = \tilde{\mathbf{W}}_S + \mathbf{e}_S$, with $\tilde{\mathbf{W}}_S$ representing the deterministic part of equation (4.3) and $\mathbf{e}_S$ representing the noise vector. Similarly $\mathbf{T} = \tilde{\mathbf{W}}_T + \mathbf{e}_T$.

In this example we will use rigid body spatial transformations defined by:

$$f_{\theta,\mathbf{t}}(\mathbf{x}) = \mathbf{R}_\theta(\mathbf{x} - \mathbf{c}) + \mathbf{c} + \mathbf{t}, \tag{4.13}$$

where $\mathbf{t} \in \mathbb{R}^d$ represents a translation vector, $\mathbf{c}$ represents the center coordinate of the image, and

$$\mathbf{R}_\theta = \begin{vmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{vmatrix}, \tag{4.14}$$

represents a rotation matrix, in the two dimensional case. For shorthand notation we write $\mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}$ to mean the operation of applying a spatial transformation $f_{\theta,\mathbf{t}}$ to digital image $S(\mathbf{i})$ through (4.10). Note that since (4.10) is a linear operation on the sampled image values, $\mathbf{F}_{\theta,\mathbf{t}}$ is a linear operator, though in general it is not

shift invariant.

We look at the discrete $L_2$-norm similarity measure, defined by $\|\mathbf{a}\|^2 = \langle \mathbf{a}, \mathbf{a} \rangle = \frac{1}{N} \sum_{i=1}^{N} a_i^2$, of the difference between the source and target images being registered as a function of the transformation parameters:

$$\begin{aligned} I(\theta, \mathbf{t}) = \|\mathbf{F}_{\theta,\mathbf{t}}\mathbf{S} - \mathbf{T}\|^2 &= \langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{S} - \mathbf{T}, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{S} - \mathbf{T} \rangle \\ &= \langle \mathbf{F}_{\theta,\mathbf{t}}(\tilde{\mathbf{W}}_S + \mathbf{e}_S) - \mathbf{T}, \mathbf{F}_{\theta,\mathbf{t}}(\tilde{\mathbf{W}}_S + \mathbf{e}_S) - \mathbf{T} \rangle. \end{aligned}$$

Expanding all terms further we write:

$$I(\theta, \mathbf{t}) = Q_1(\theta, \mathbf{t}) + Q_2(\theta, \mathbf{t}), \tag{4.15}$$

where

$$\begin{aligned} Q_1(\theta, \mathbf{t}) &= \langle \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S, \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S \rangle + 2\langle \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S \rangle \\ &\quad -2\langle \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S, \mathbf{T} \rangle - 2\langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{T} \rangle + \langle \mathbf{T}, \mathbf{T} \rangle \end{aligned} \tag{4.16}$$

and

$$Q_2(\theta, \mathbf{t}) = \langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S \rangle. \tag{4.17}$$

To illustrate the behavior of $Q_1(\theta, \mathbf{t})$ and $Q_2(\theta, \mathbf{t})$ in the presence of noise we have computed the following simulation example. Normally distributed spatially uncorrelated noise was added to a digitally manufactured 'phantom' image (shown in figure 4.3) such that the signal to noise ratio, defined to be mean signal divided by the standard deviation, was about 22. Two such images were generated with identical signal and different noise vectors with one considered as the source image and the other the target image. Here optimal alignment is defined to be at zero degrees of rotation and zero pixel translations.

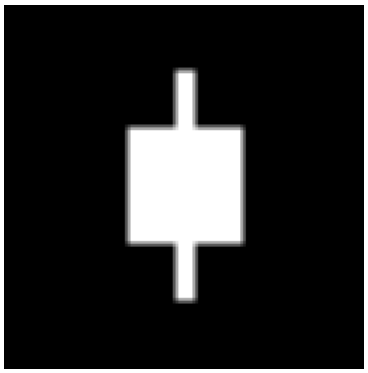

Figure 4.3: Synthetic digital phantom image used in simulations. Shown here is only the deterministic part (no noise).

$Q_1(\theta, \mathbf{t})$ and $Q_2(\theta, \mathbf{t})$ were then computed first as a function of image rotation (figure 4.4) and then with respect to translation along the x direction (figure 4.5) by using bilinear interpolation. Evidently $Q_1(\theta, \mathbf{t})$ is able to determine the transformation parameters that optimally align the images. Inspecting equation (4.16) more closely we see that $\langle \mathbf{F}_{\theta,\mathbf{t}} \tilde{\mathbf{W}}_S, \mathbf{F}_{\theta,\mathbf{t}} \mathbf{e}_S \rangle \sim 0$ since, by definition, $\mathrm{E}\{e_S(\mathbf{x})\} = 0 \; \forall \; \mathbf{x}$. By a similar argument $\langle \mathbf{F}_{\theta,\mathbf{t}} \mathbf{e}_S, \mathbf{T} \rangle \sim 0$. The remaining terms in $Q_1$ can be written as $\|\mathbf{F}_{\theta,\mathbf{t}} \tilde{\mathbf{W}}_S - \tilde{\mathbf{W}}_T\|^2 + \langle \mathbf{e}_T, \mathbf{e}_T \rangle$, where $\mathrm{E}\{e_T(\mathbf{x})\} = 0 \; \forall \; \mathbf{x}$ is used again. Since no significant terms involving $\langle \mathbf{F}_{\theta,\mathbf{t}} \mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}} \mathbf{e}_S \rangle$ appear, we see why $Q_1$ is, in theory, artifact free. $Q_2(\theta, \mathbf{t})$, on the other hand, oscillates with respect to the spatial transformation being applied. Because no terms containing the signal part of the images is present the oscillations in $Q_2(\theta, \mathbf{t})$ are independent on the alignment of image signal, thus creating a confound in the optimization of the objective function. As a consequence, $L_2$-based similarity measures $I(\theta, \mathbf{t}) = \|\mathbf{F}_{\theta,\mathbf{t}} \mathbf{S} - \mathbf{T}\|^2 = Q_1(\theta, \mathbf{t}) + Q_2(\theta, \mathbf{t})$ are not able to determine the transformation parameters that optimally align the images when significant noise is present. Thus a computer program that registers images based on the minimization of equation $\|\mathbf{F}_{\theta,\mathbf{t}} \mathbf{S} - \mathbf{T}\|^2$ would not be able to produce accurate

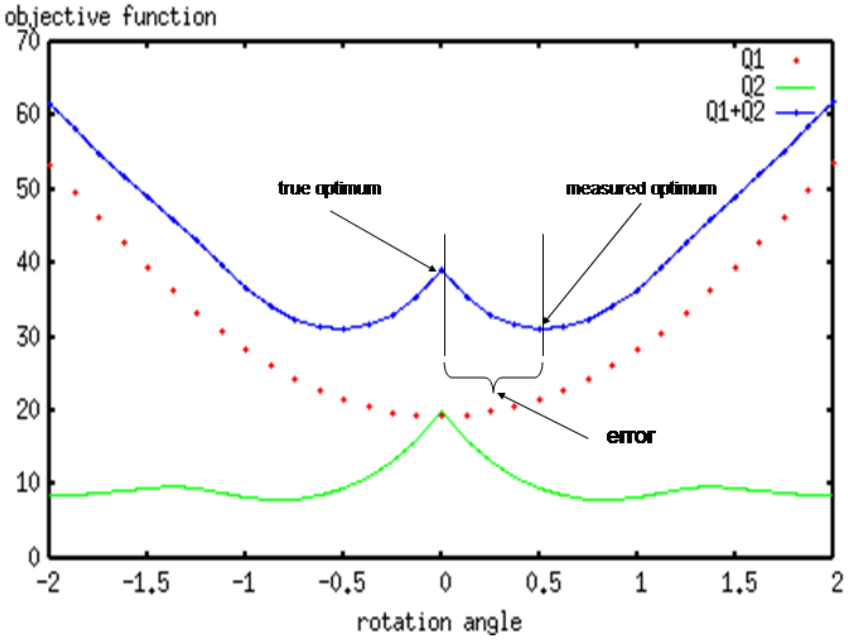


Figure 4.4: Results from simulation experiments. Sum of squared differences similarity measure with respect to rotation angle (in degrees) computed using bilinear image interpolation. The sum of squared differences (dotted line) is computed as a sum of two terms: $Q_1(\theta)$ and $Q_2(\theta)$ (see text). Zero degrees rotation defines optimal alignment. True optimal alignment and measured optimal alignment are indicated with arrows. Magnitude of the error in the computed registration parameter in this case is 0.5 degrees.

results. Figures 4.4 and 4.5 show the size of the error in the estimated registration parameter such a program is likely to produce in this specific example. The error in rotation would be about $\pm 0.5$ degrees while the error in translation would be about $\pm 0.25$ pixels.

Looking at $Q_2(\theta, \mathbf{t})$ more closely, we see that, when $e_S(\mathbf{x})$ has a normal distribution (other distributions may also apply), $\langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S \rangle$ is the maximum likelihood estimator for the expectation $\mathrm{E}\{(\mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S)^2\}$. For the case of image trans-

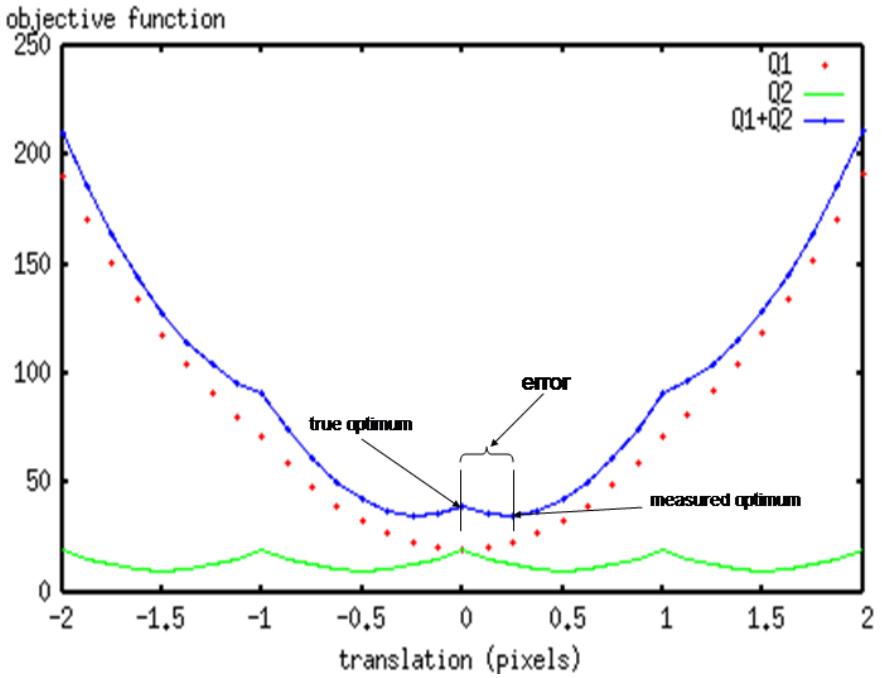


Figure 4.5: Results from simulation experiments. Sum of squared differences similarity measure with respect to translation (in pixels) computed using bilinear image interpolation. The sum of squared differences (dotted line) is computed as a sum of two terms: $Q_1(\theta)$ and $Q_2(\theta)$ (see text). Zero translation defines optimal alignment. True optimal alignment and measured optimal alignment are indicated with arrows. Magnitude of the error in the computed registration parameter in this case is 0.25 pixels.

lation, it is easy to see that $\mathrm{E}\{(\mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S)^2\} \sim \mathrm{Var}\{S^c(f(\mathbf{x}))\}$, where $\mathrm{Var}\{S^c(f(\mathbf{x}))\}$ is given by equation (4.9). The oscillations in $Q_2(\theta, \mathbf{t})$ shown in figure 4.5 are thus equal (up to a scaling factor) to the oscillations shown in the linear portion of figure 4.1. Note also that the oscillations shown in figure 4.5 are very similar to the 'grid' effects discussed in [85, 112].

Finally, we note that the constant variance assumption is not strictly necessary to explain the so called 'grid' artifacts in the sum of squared differences similarity

measure, nor should it be expected that such artifacts will only occur in which the constant variance assumption holds. To see this, consider the $Q_2$ (equation 4.17) term in the expanded sum of squared differences similarity measure:

$$Q_2(\theta, \mathbf{t}) = \langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S \rangle = \frac{1}{N}\sum_{i=1}^{N} \left(\mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S\right)_i^2 . \tag{4.18}$$

For noise that has an exponential type of distribution (Rayleigh, Gaussian, etc.), and in the case of image translation specifically, the sum above can easily be shown to be the maximum likelihood estimator of the variance at any given interpolated pixel. Now, the sum above can be divided into sub-sums over regions where it can be assumed that, at least locally, the constant variance assumption holds. This is appropriate for the case of MRI for example, where several researchers have shown that the noise distribution in the background satisfies a Rayleigh-type distribution (assuming no significant ghosting contamination) of constant variance [43, 100]. In the foreground, at signal to noise ratios greater than 3 or so, the noise distribution approximates a Gaussian distribution, also of constant variance [49, 43, 100]. Now the sum above can be split into sub-sums, where the constant variance assumption holds. Thus

$$\frac{1}{N}\sum_{i=1}^{N} \left(\mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S\right)_i^2 = \frac{1}{N}\left(\sum_{i\in\Omega_1} \left(\mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S\right)_i^2 + \sum_{i\in\Omega_2} \left(\mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S\right)_i^2\right), \tag{4.19}$$

where $\Omega_1$ and $\Omega_2$ represent two regions where the constant variance assumption holds. Naturally, the same argument can be expanded for and arbitrary number of regions. It is easy to see that each individual term $\sum_{i\in\Omega_j} \left(\mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S\right)_i^2$ is proportional to the variance formula 4.9, which as explained earlier, oscillates with respect to the spatial transformation being applied to the image.

### 4.2.3 Optimization of correlation-based similarity measures

In this section we look at the effects of system noise on correlation based similarity measures. For convenience, we look at the cross correlation (zero-mean correlation coefficient) between two $N$-dimensional vectors often used in image registration:

$$I(\theta, \mathbf{t}) = \frac{\langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}, \mathbf{T} \rangle}{\|\mathbf{T}\|\|\mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}\|}. \tag{4.20}$$

As done earlier, we use the linear stochastic image model $\mathbf{S} = \tilde{\mathbf{W}}_S + \mathbf{e}_S$ to expand the term $\|\mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}\|$ into

$$\begin{aligned} \|\mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}\| &= \sqrt{\langle \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S + \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S + \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S \rangle} \\ &= \sqrt{\langle \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S, \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S \rangle + 2\langle \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S \rangle + \langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S \rangle} \end{aligned}$$

where again we have the undesirable, though inevitable, term $\langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S \rangle$ which depends solely on the noise properties of the source image. The same translation and rotation simulations reveal that correlation based cost functions suffer from the same problems as $L_2$ based ones: namely, the optimal value of the cost function will not reveal the transformation parameters that optimally align the images. The results of the simulation are shown in figure 4.6. Again, the assumption of globally constant noise variance is not strictly necessary, since the term $\langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S \rangle$ can be split over regions where it can be assumed the variance is constant, by a similar argument to the one presented in the previous section.

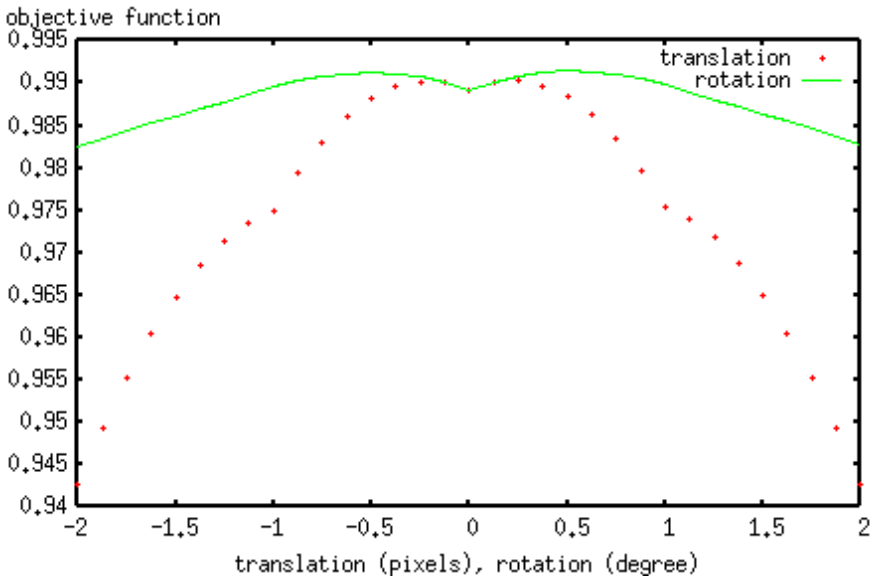


Figure 4.6: Results from simulation experiments. Cross correlation similarity measure as a function of image translation (dots) and rotation (solid).

### 4.2.4 Optimization of mutual information

The 'grid' effects observed in mutual information-based optimization are slightly different in nature from the effects demonstrated for the $L_2$ and correlation objective functions in the sense that noise variance is not the only factor to be considered. In this case, we have also to consider the intensity value distribution of the object, in addition to the distribution of noise. Image registration via maximization of mutual information relies on measuring the statistical dependence in the co-occurrence of intensity values of images $S(f(\mathbf{x}))$ and $T(\mathbf{x})$. Let $\mathrm{pr}_{S_f}[s]$ and $\mathrm{pr}_T[t]$ represent the distribution of pixel intensity values $s$ and $t$ in images $S(f(\mathbf{x}))$ and $T(\mathbf{x})$, respectively. Their joint probability is denoted $\mathrm{pr}_{S_f,T}[s,t]$. With these quantities defined the mutual information between images $S(f(\mathbf{x}))$ and $T(\mathbf{x})$ is given by:

$$I\left(S(f(\mathbf{x})), T(\mathbf{x})\right) = \int\int \mathrm{pr}_{S_f,T}[s,t] \log\left(\frac{\mathrm{pr}_{S_f,T}[s,t]}{\mathrm{pr}_{S_f}[s]\mathrm{pr}_T[t]}\right) dsdt, \tag{4.21}$$

or alternatively

$$I\left(S(f(\mathbf{x})), T(\mathbf{x})\right) = H(T) + H(S_f) - H(T, S_f), \tag{4.22}$$

where

$$\begin{aligned} H(T) &= -\mathrm{E}\{\log\left(\mathrm{pr}_T[t]\right)\}, \\ H(S_f) &= -\mathrm{E}\{\log\Big(\mathrm{pr}_{S_f}[s]\Big)\}, \\ \text{and } H(S_f, T) &= -\mathrm{E}\{\log\Big(\mathrm{pr}_{S_f,T}[s,t]\Big)\}. \end{aligned}$$

As an initial example, let $\mathrm{pr}_T[t] = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left(-\frac{(t-\mu_1)^2}{2\sigma_1^2}\right)$ and similarly $\mathrm{pr}_{S_f}[t] = \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left(-\frac{(t-\mu_2)^2}{2\sigma_2^2}\right)$, where $\sigma_1^2$ and $\sigma_2^2$ are the variances of the distribution of intensity values of images $S(f(\mathbf{x}))$ and $T(\mathbf{x})$, respectively, and $\mu_1$, $\mu_2$ their means. Then the mutual information between images $S(f(\mathbf{x}))$ and $T(\mathbf{x})$ is given by [61]:

$$I\left(S(f(\mathbf{x})), T(\mathbf{x})\right) = -\frac{1}{2}\log\left(1-\rho^2\right), \tag{4.23}$$

where $\rho$ is the linear correlation coefficient between the intensity values of the images being registered. In this situation we would expect that any reasonably accurate estimate of (4.21) as a function of spatial transformation $f$ would also contain the artifacts in the correlation coefficient similarity measure demonstrated earlier. Naturally, tomographic images are seldom globally Gaussian distributed. Locally, however, the normal distribution assumption may be more realistic. In appendix C of this chapter we explain that such oscillations occur for a mixture

model specific to MRI at high signal to noise ratios. While this is not an analytical explanation for all possible distributions, for all possible types of images, we believe it covers at least one realistic circumstance (that of MRI at high SNR) as well as other similar cases.

To show the effects of low order interpolation on mutual information based optimization of spatial transformations we have computed the same translation simulation described earlier. We have used a Parzen windowing technique to compute an estimate of the joint probability density function $\mathrm{pr}_{S_f,T}[s,t]$ based on discretely sampled values of the images $S(f(\mathbf{x}))$ and $T(\mathbf{x})$. That is

$$\mathrm{pr}_{S_f,T}[s,t] \sim \sum_{i=1}^{N} \psi\left(S(f(\mathbf{x}_i)) - s\right) \psi\left(T(\mathbf{x}_i) - t\right), \tag{4.24}$$

where $\psi(x) = \frac{1}{\beta\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\beta^2}\right)$, and $\mathbf{x}_i,\ i = 1, \cdots, N$ are the sampling coordinates of the target image $T(\mathbf{x})$. The results of the simulation are shown in figure 4.7. Evidently, even for this simple simulation, the mutual information similarity measure as a function of image translation contains the multiple local optima 'grid' artifacts described earlier.

## 4.3 Methods

The theory presented above suggests that the interpolation artifacts seen in $L_2$ and correlation-based cost functions are entirely due to the presence of noise in the images being registered. One obvious strategy to mitigate such artifacts is to reduce the variance $\mathrm{Var}\{S(\mathbf{i})\}$ in the original source image by performing local averaging. This can be achieved via digital convolution of the image with a digital filter $G(\mathbf{i})$:

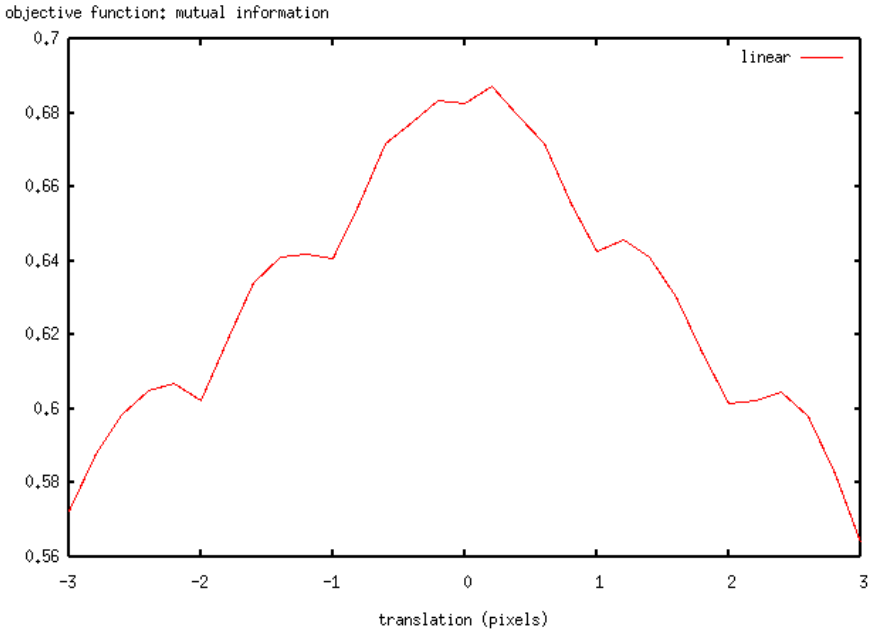


Figure 4.7: Results from simulation experiments. Mutual information as a function of image translation computed using bilinear interpolation.

$$\tilde{S}(\mathbf{i}) = \sum_{\mathbf{w}\in\mathbb{Z}^d} S(\mathbf{w})G(\mathbf{i}-\mathbf{w}). \tag{4.25}$$

If we are concerned with a stationary random process, $\mathrm{Var}\{S(\mathbf{i})\} = \sigma^2 \ \forall \ \mathbf{i}$, using

$$G(\mathbf{x}) = \frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \tag{4.26}$$

would reduce the variance of the image to $\mathrm{Var}\{\tilde{S}(\mathbf{i})\} = 0.11\sigma^2$. This would reduce the oscillatory behavior of the cost function due to the term $\langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S \rangle$ in computing $L_2$ and correlation based cost functions, and may also reduce the oscillatory behavior of mutual information curves [85] . This may not always be an optimal strategy for excessive smoothing can blur image boundaries that are important for guiding the registration process. As suggested earlier, an alterna-

tive approach is to use higher order sinc approximating kernels instead of low order kernels such as the hat function.

We will compare these interpolation and approximation methods for measuring image similarity in real and simulated magnetic resonance images. The simulated phantom image is shown in figure 4.3. Uncorrelated normally distributed noise was added in quadrature (in MRI time domain signals are usually received in quadrature and the displayed image is usually the magnitude of the inverse Fourier transform of the received signal [44]) such that the signal to noise ratio was about 31. The real MR images shown in figure 4.8 were taken from a standard single-shot spin-echo planar imaging (EPI) sequence acquisition on a 1.5 T GE Signa system. Because the images were reconstructed by taking the magnitude of the Fourier transform of the time domain signals, they do not necessarily constitute a stationary random process. In fact through the work described in [49, 43] it is known that if Gaussian distributed noise is added to the receiver coils, the variance of the magnitude reconstructed image in regions of zero signal is $(2-\pi/2)\sigma^2$ while the variance of the magnitude reconstructed image in regions with relatively high signal approaches $\sigma^2$. The variance in each of these domains however, is usually assumed to be constant. In addition, because of several linear filtering steps, performed during analog to digital conversion, to remove 'ringing' artifacts, the magnitude reconstructed is usually spatially correlated. However, these correlations are usually small.

In addition to the echo-planar $T_2$ images shown in figure 4.8, we have also used images taken from a realistic MRI simulator [24]. The image pair consists of noisy high resolution $T_2$ and $T_1$ weighted images. Computation with these images are included to make explicit the fact that our theoretical predictions are

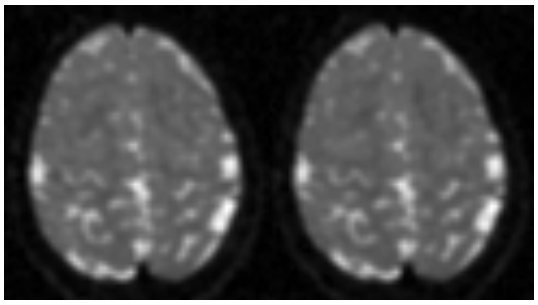

Figure 4.8: Real $T_2$ weighted echo planar images used for computing the performance of different interpolation methods for image registration. The images were acquired in rapid succession and are likely to be fairly well aligned.

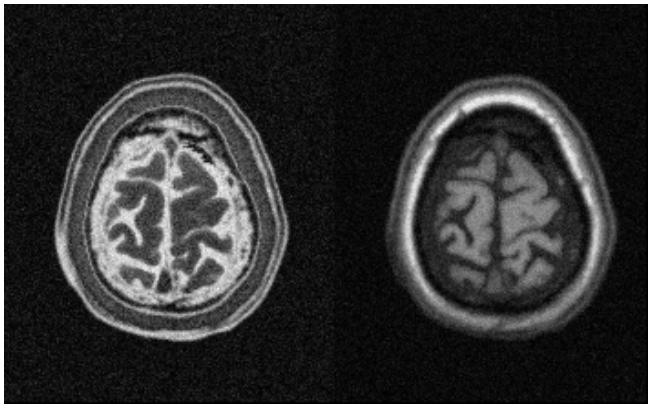

Figure 4.9: Images used for computation of mutual information similarity measure.

not limited to the data set shown in figure 4.8.

As done in [85, 112] the registration curves were computed by rotating and translating one image with respect to the other, much like in the previously discussed simulations. However, following the approach described in [117], in order to amplify the interpolation artifacts we choose to vary the spatial transformation starting from a slightly misaligned state. That is, angle $\theta$ is varied after translating the source image using a fixed translation vector $\mathbf{t} = \{0, 3\}^T$. Similarly, translation in the $x$ direction is computed after translating the source image

using a fixed translation in the y direction $\mathbf{t} = \{t, 3\}^T$. The transformations are combined so that only one image interpolation is necessary for each value in the computed registration curves. Note that when real data was used, the exact alignment parameters are unknown. However, because the images were acquired in rapid succession they are likely to be closely aligned. This can be confirmed by visual inspection of the images displayed in figure 4.8.

In order to show that the effects described above are general and not limited to affine or 'global' transformations we have also conducted tests using spatially varying transformations constructed with localized radial basis functions. In this example we used the following parameterization for the spatial transformation:

$$f_{\mathbf{k}}(\mathbf{x}) = \mathbf{x} + \mathbf{k}\Phi(\frac{\|\mathbf{x} - \mathbf{q}\|}{r}), \tag{4.27}$$

where $\mathbf{k} = \{k_x, k_y\}^T$ are the basis functions coefficients, $\mathbf{q}$ defines the center of the basis function, $r$ its radius, and:

$$\Phi(x) = \frac{1}{4}(1-x)_+^4(4 + 16x + 12x^2 + 3x^3). \tag{4.28}$$

As in the example described earlier, we have plotted the variation of the cost function with respect to the $k_x$ coefficient, using a fixed $k_y = 2$ coefficient.

Finally, we point out that in many useful cases spatial transformations using sinc-type basis functions can be computed efficiently using the FFT algorithm, following the arguments highlighted in [31]. Let $S(n),\ n = 0, \cdots, N-1$ represent a one dimensional discrete vector, the same concepts can be easily generalized to multiple dimensions. The Discrete Fourier Transform (DFT) of the signal is:

$$\hat{S}(k) = \sum_{n=0}^{N-1} S(n) e^{-\frac{j2\pi kn}{N}} \tag{4.29}$$

As shown in Eddy *et al* [31], a pure translation of the image vector by $t \in \mathbb{R}$ can be obtained by multiplying $\hat{S}(k)$ with a complex exponential followed by inverse discrete Fourier transformation:

$$S(n-t) = \mathcal{F}^{-1}\left(e^{-\jmath 2\pi tk}\hat{S}(k)\right). \tag{4.30}$$

In image domain, the above operation is equivalent to performing a discrete convolution, with periodic boundaries, of the original image vector with the Dirichlet kernel

$$h(x) = \frac{\sin(\pi x)}{N\sin\left(\pi/Nx\right)}\exp\left(-\jmath\pi x(N-1)/N\right) \tag{4.31}$$

sampled at $x = t - n,\ n = 0, \cdots, N-1$. As $N \to \infty$ the kernel above converges to the already mentioned sinc kernel. It is easy to see that signal translation via the DFT algorithm described above is also optimal in the sense that it does not corrupt the power spectral density of a stationary random process [13] since $|e^{-\jmath 2\pi tk}|^2 = 1\ \forall\ tk$. We have implemented such image translation algorithm to compare to other translation methods based in linear and cubic interpolation.

## 4.4 Results

Figure 4.10 shows the sum of squared differences cost function with respect to translation of the digital phantom image described above. The figure shows that bilinear interpolation is not an appropriate interpolation procedure for performing image registration when significant noise is present. In the location where a global minimizer is expected, bilinear interpolation-based computation of the objective function seems to produce a local maximizer. Sinc-based interpolation, here the width of the windowing function was $W = 6$, however, seems to perform well. Similar results are obtained using rotations, instead of translations. Similar

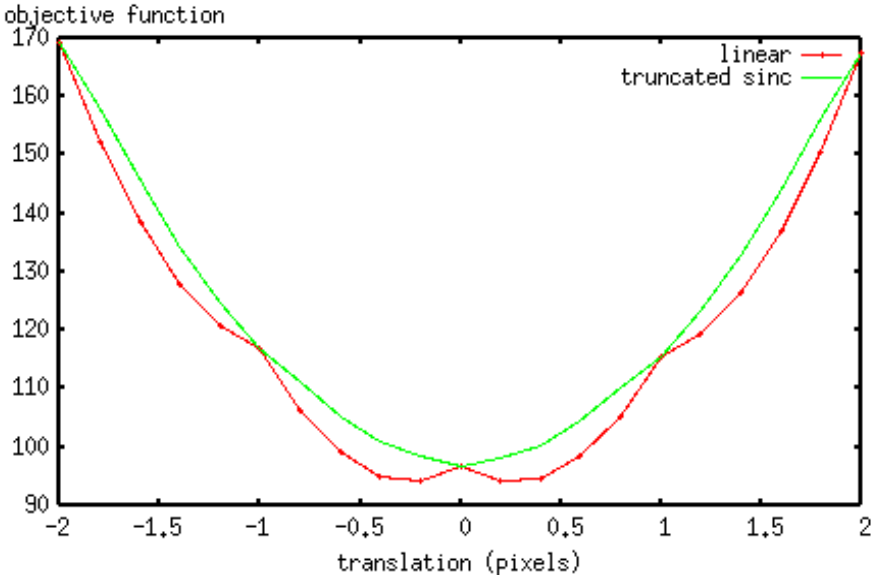


Figure 4.10: Sum of squared differences as a function of image translation for digital phantom image experiments. The dotted line curve was computed using bilinear interpolation while the solid curve was computed using truncated sinc interpolation.

results are obtained using correlation and mutual information similarity measures [90]. For brevity these are omitted here, and we now focus on the experiments using real MR images.

Figure 4.11 shows the sum of squared differences similarity measure (using the real images displayed in figure 4.8) with respect to translation using the bilinear and truncated sinc interpolation methods. The registration curve computed using bilinear interpolation presents what is commonly refereed to as the grid artifact while such artifacts are seemingly non existent in the registration curve computed using the truncated sinc interpolation method. Figure 4.12 shows similar results computed using image rotation instead of translation. Both figures 4.11 and 4.12 were computed using $W = 6$ for the truncated sinc basis function. Note

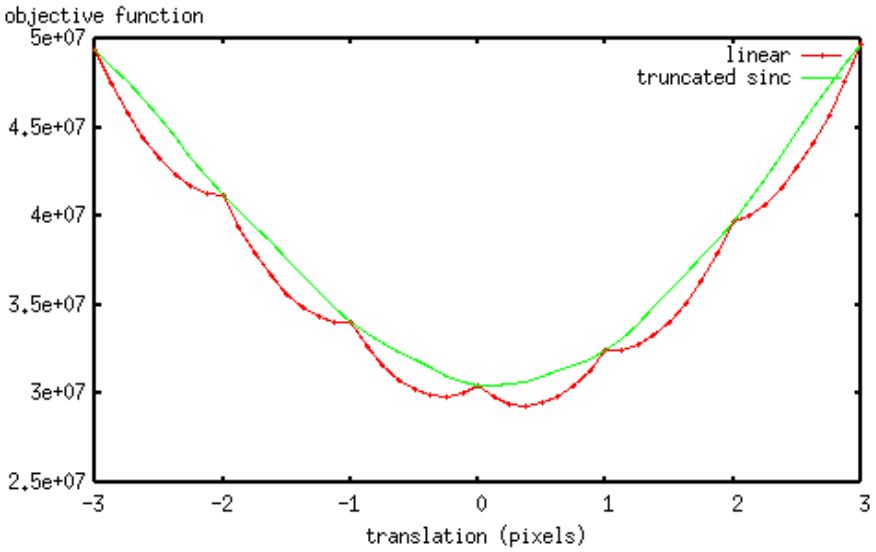


Figure 4.11: Sum of squared differences as a function of image translation for real MR data experiments. The dotted line curve was computed using bilinear interpolation while the solid curve was computed using sinc approximating basis functions.

that in both cases, bilinear interpolation produces different global optima with respect to transformation parameters than truncated sinc interpolation. Figure 4.13 displays again the sum of squared differences similarity measure computed with respect to image translation. In this example, however, the source image was blurred by the convolution filter defined in equation (4.26). The grid artifact is substantially reduced in comparison to figure 4.11, though it is not completely eliminated.

Figures 4.14,4.15, and 4.16 display the results of the same experiments using the cross correlation objective function. In this example truncated sinc interpolation was computed using $W = 12$. Again the grid artifact is fairly evident in the translation curves computed using bilinear interpolation (figure 4.14). The grid artifacts are removed by using truncated sinc interpolation. The rotation exper-

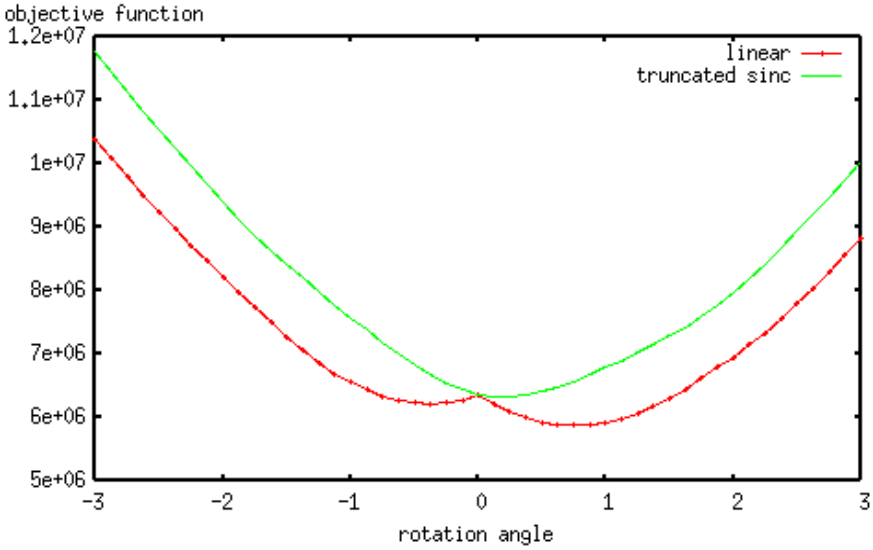


Figure 4.12: Sum if squared differences as a function of image rotation for real MR data experiments. The dotted curve was computed using bilinear interpolation while the solid curve was computed using truncated sinc basis functions.

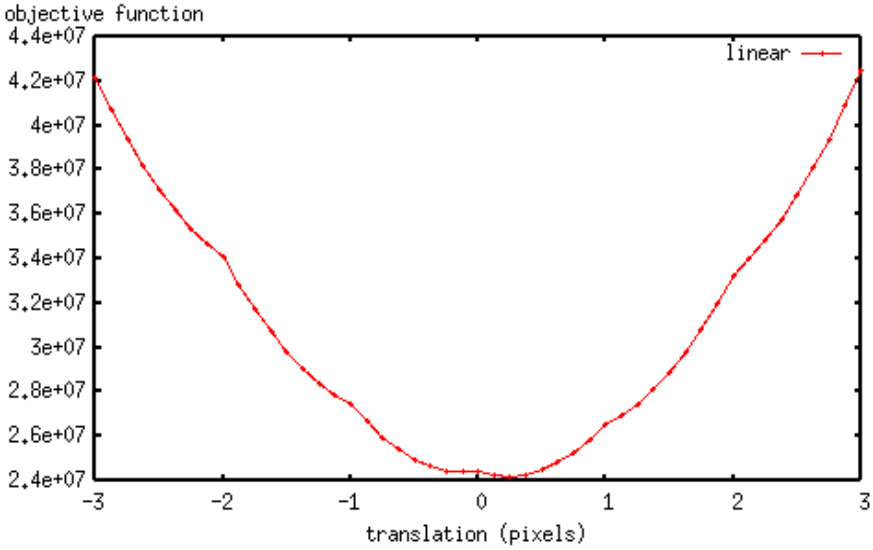


Figure 4.13: Sum of squared differences similarity measure as a function of image translation for real data experiments. In this experiment the image being rotated was blurred prior to computation of the similarity measure.

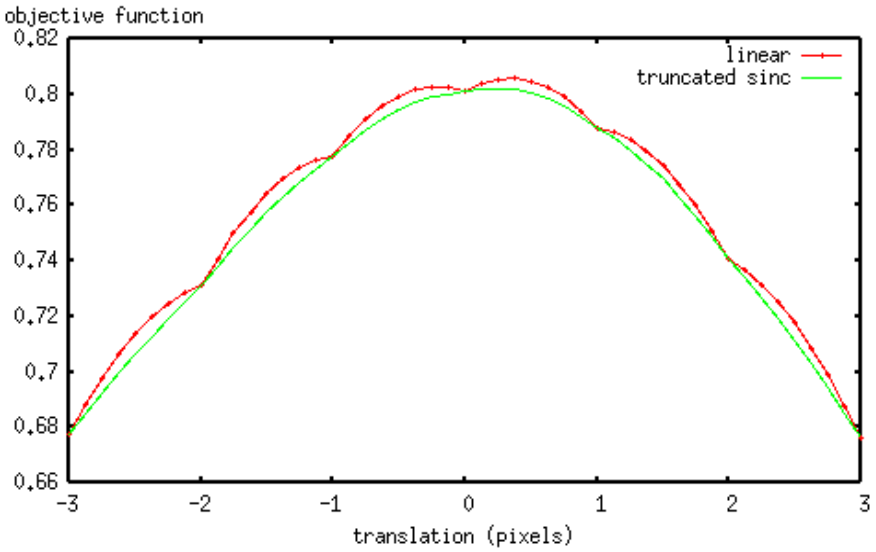


Figure 4.14: Cross correlation as a function of image translation for real data experiments. Solid curve computed using truncated sinc basis functions while the dotted curve was computed using bilinear interpolation.

iments reveal that one encounters a local minimum in the objective function at zero degrees (figure 4.15). The local optimum disappears when using truncated sinc. Once again, the grid effects seen to be reduced when using blurred images (figure 4.16).

Figures 4.17,4.18,4.19 display the results of the experiments performed using the mutual information similarity measure. Here the width of the truncated sinc interpolation kernel was $W = 50$. Translation experiments reveal that linear interpolation causes the aforementioned grid artifacts, while truncated sinc interpolation seems to avoid them. The rotation experiments in this example were less clear than others. However, the oscillatory behavior of the registration curve is much reduced when using sinc approximating interpolation kernels. As with the previous experiments, smoothing of the source image prior to interpolation and similarity measure computation seems to ease grid artifacts.

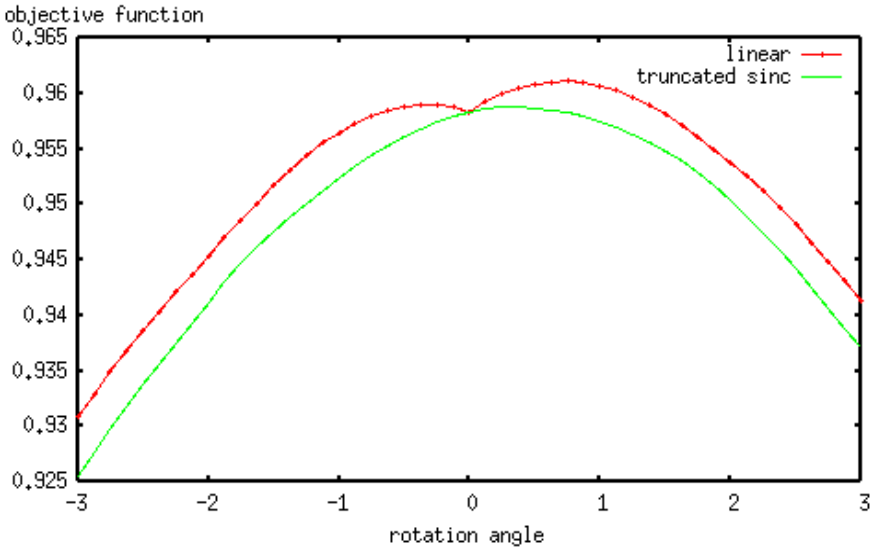


Figure 4.15: Cross correlation as a function of image rotation for real data experiments. Solid curve computed using truncated sinc basis functions while the dotted curve was computed using bilinear interpolation.

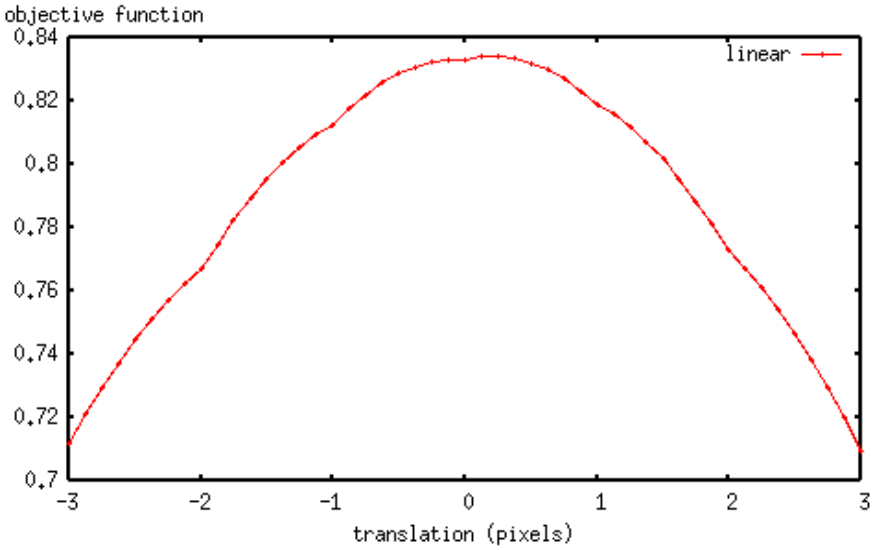


Figure 4.16: Cross correlation as a function of image translation. In this example the source image was blurred prior to computation of the similarity measure.

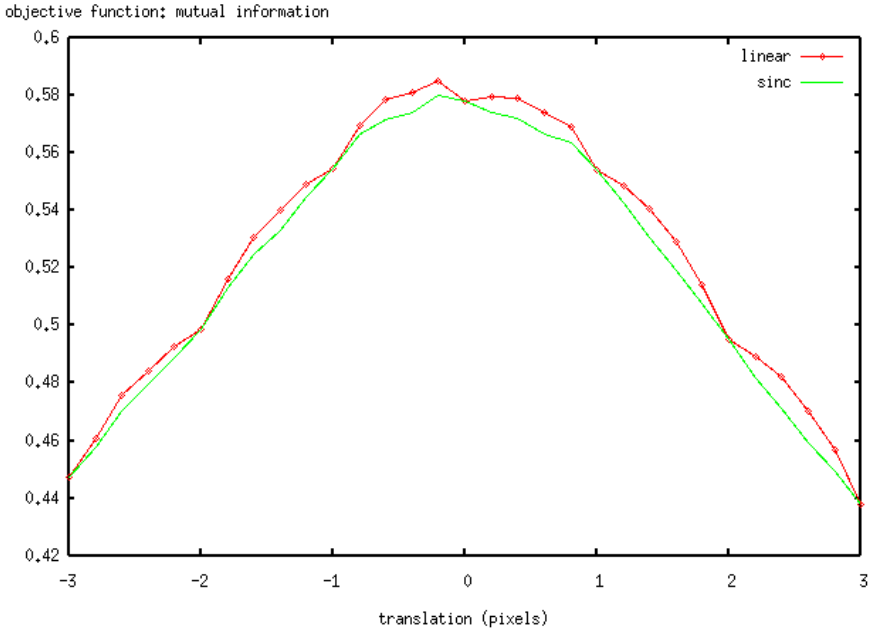


Figure 4.17: Mutual information as a function of image translation for real image experiments. Solid curve computed using truncated sinc basis function while the dotted curve was computed using bilinear interpolation.

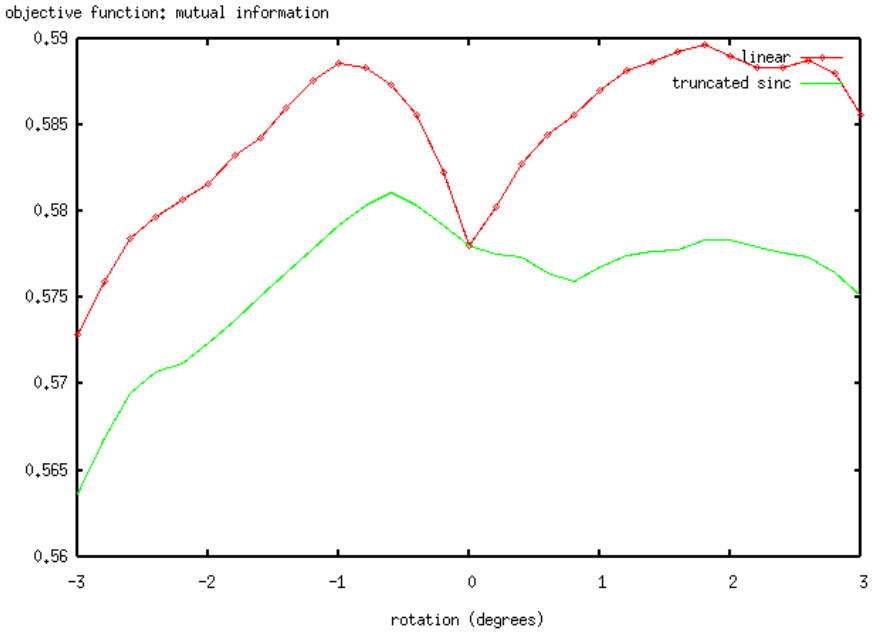


Figure 4.18: Mutual information as a function of image rotation (in degrees) for real image experiments. Solid curve computed using truncated sinc basis function while the dotted curve was computed using bilinear interpolation.

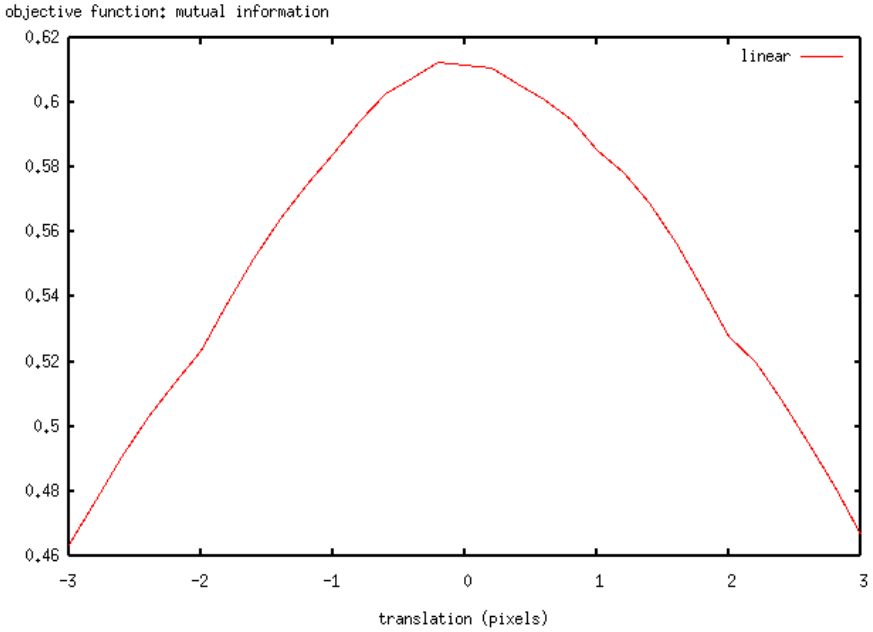


Figure 4.19: Mutual information as a function of image translation for real image experiments. Computed using bilinear interpolation after smoothing the source image.

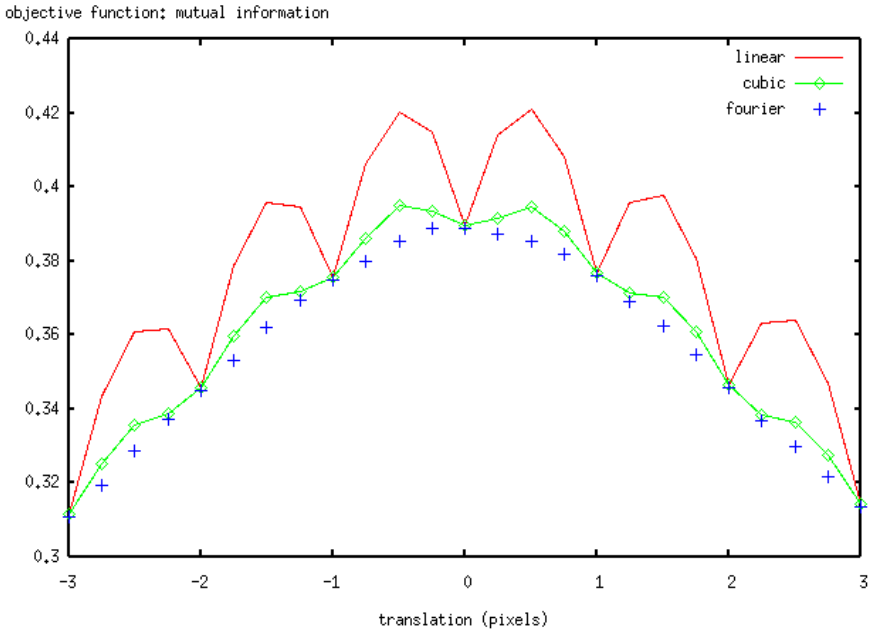


Figure 4.20: Mutual information similarity measure with respect to image translation computed using linear, cubic, and fourier-based interpolation.

Figure 4.20 contains plots of the mutual information similarity measure with respect to translation for the images shown in figure 4.9. In this specific example, the *pdf*s used to estimate the mutual information were computed using the joint histogram of the images, in a manner similar to the method described in [70]. Three registration curves were computed. One using linear interpolation, one using cubic interpolation, and the other using the FFT-based translation algorithm described earlier. As evident from these plots, the FFT-based image translation method seems to be the best choice for determining the translation parameters that register the images. In addition, observe that the cubic interpolation kernel seemed superior to the linear one in the sense that the oscillations in the registration curve was greatly reduced.

Finally, the results of the nonrigid experiments using the $L_2$-norm objective function are displayed in figure 4.22. Here truncated sinc interpolation was com-

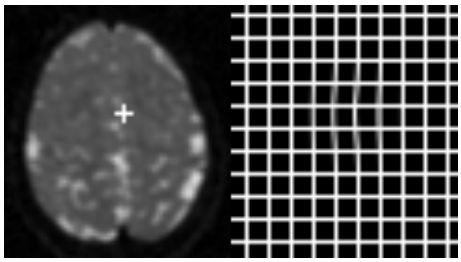

Figure 4.21: Left: MR image used for local, nonrigid, deformation experiment. The white cross close to the center of the image indicates the location of the basis function. Right: local deformation computed using the compactly supported basis function. The location coinciding with the center of the basis function is moved one pixel to the right.

puted using $W = 6$. The location of the basis function is indicated by the bright cross in the image shown in the left panel figure 4.21. The radius of the basis function was chosen to be $r = 20$ pixels. The deformation produced by translating the center image location one pixel to the left is shown visually on the right panel of the same figure. The curve computed using bilinear interpolation again shows multiple local optima values. In contrast, the curve computed using truncated sinc interpolation shows only one. In addition, the global optimum for the cost function computed using truncated sinc basis functions and bilinear ones differ.

## 4.5 Discussion

The experiments with real MR images discussed above depict the effects of low order interpolation methods on popular registration cost functions such as the sum of squared differences, the correlation coefficient, and mutual information. Though the optimal alignment parameters for the real data experiments were not

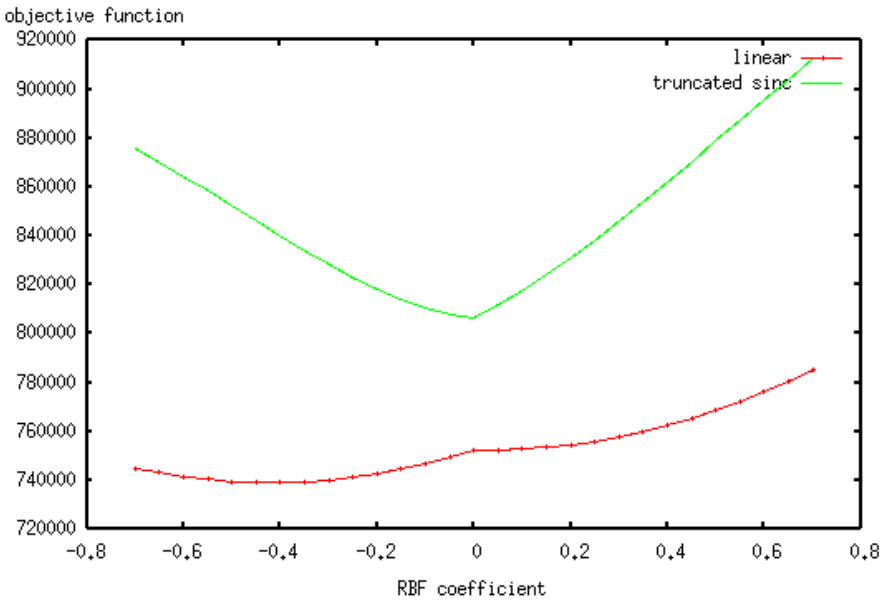


Figure 4.22: Sum of squared differences similarity measure as a function of local deformation using a compactly supported radially symmetric basis function. The solid curve was computed using truncated sinc interpolation while the dotted curve was computed using bilinear interpolation.

known, in all experiments tried, sinc approximating basis functions performed visibly better than linear ones. Linear interpolation often produced registration curves that contained the aforementioned grid (multiple local optima) artifacts. Thus registration programs based on optimization of $L_2$, linear correlation, or mutual information-based cost functions using linear interpolation are likely to converge slowly to the global optima solution, or in some cases, converge only to a local optimum, significantly degrading the quality of the results. Image interpolation using sinc approximating basis functions, on the other hand, did not produce the artifacts mentioned.

Local parametric models for deformation fields are often used in nonrigid registration methods [95, 64, 65, 87]. The results produced by using the local nonrigid transformation model (4.27) show that the interpolation artifacts in registration curves are general, and not necessarily restricted to affine or global transformation models. Results also showed that sinc approximating kernels can help increase the accuracy of such registration algorithms.

The FFT-based algorithm seems to be a good alternative for computing linear-type spatial transformations of digital images. The algorithm explained earlier in this chapter was specific for computing pure translations. However, it can be generalized to affine transformations, as shown by Eddy *et al.* [31]. As explained earlier, Fourier interpolation via the FFT is not only computationally efficient, but also has the benefit of not distorting the stochastic properties (variance and covariance) of the image being transformed. Thus similarity measures computed using the FFT image transformation algorithm are more likely to be free of the so called 'grid' artifacts (local optima). This is a powerful alternative to truncated sinc kernel which, in order to approximate the properties of true sinc

interpolation to the fidelity necessary to remove 'grid' artifacts, require extremely long computation times.

We attribute the increase in performance gained by sinc approximating kernels to the fact that their use in computing geometric image transformations causes the least amount of changes in the covariance properties of the image being interpolated. This was particularly evident when we compared sinc approximating basis functions with linear basis functions in our simulations. We have experimented with different apodisation windows for truncating sinc basis functions and found that the support $W$ of the windowing function, and not necessarily the form of the windowing function, was most significant in determining the covariance properties of interpolated images [90].

Finally, it is worth pointing out that the theory and experimental results presented here are in good agreement with the more empirical findings published by many elsewhere. Blurring the data has been shown to reduce local optima artifacts by Woods *et al* [122] and Ashburner *et al* [9], while W.F. Eddy *et al* [31] and J.V. Hanjal *et al* [46] have described advantages of sinc type interpolation methods in the context of image registration. Shift invariant image blurring was apparently not used by Pluim *et al* [85], though a "resampling" operation was reported to reduce the severity of artifacts. Such resampling will reduce the variance of each local sample. As pointed out in the paper, reduction in the variance of the image will help reduce such artifacts. Finally Tsao [112] reported that a sinc based kernel of width $W = 3$ was not effective in diminishing artifacts in mutual information curves, something we also experienced. However, sinc type interpolation using wider kernel widths was not reported in Tsao [112].

In fact, the only apparent contradiction is that Tsao [112] reported that higher

order interpolation such as cubic or sinc did not present any apparent advantage in the computation of MI curves. Our theory section, as well as some of the results included clearly describes advantages of higher order interpolation methods such cubic interpolation over linear type (as far as the stochastic properties of the signal are concerned). However, taking a closer look at figure 7 of Tsao's work [112], it does seem that in that specific example the cubic-type interpolation method and Hamming-sinc (even with the short support used) did reduce the 'grid' oscillations as compared to the linear interpolation method. Why this was not reflected in their quantitative measures of registration curve smoothness we do not know, and can only speculate. It could be related to things such as their preprocessing of the images, the construction of the smoothness measures, etc.

## 4.6 Summary and Conclusions

Image similarity measures based on the $L_2$ norm, linear correlation, and mutual information have been widely employed for rigid-body, affine, and elastic registration of medical and other types of images. We have shown that such similarity measures contain systematic local optima artifacts when low order interpolators such as the hat function are used for computing geometric transformations. These artifacts present a significant confound to any program that seeks to align images based on optimization of such objective functions. We showed that the interpolation artifacts are entirely caused by the effects of the spatial transformations being used as well as the basis functions used to make the images continuous.

Using a linear stochastic model for the image data we showed that in addition to being functions of the spatial transformations being applied, $L_2$ and correlation based similarity measures are also functions of the covariance structure of

the interpolated images. The covariance structure of geometrically transformed images is described as follows. Because of inevitable system noise stemming from thermal and other effects, the images being registered are viewed as random processes. Since image interpolation, at each point in space, can be interpreted as a linear filtering procedure we used the second order theory of random processes to describe the effects of different interpolating basis functions on the covariance structure of the images.

We showed that sinc approximating basis functions are optimal for generating continuous approximations of image data in registration problems since they are least likely to cause significant changes in the covariance structure of interpolated images. Thus, in addition to the more traditional properties based on linear approximation theory described in [108], we argue that the criterion $\sum_{\mathbf{i}=-\infty}^{+\infty}[h(\mathbf{x}-\mathbf{i})]^2 = 1 \; \forall \; \mathbf{x}$ should be kept in mind when choosing basis functions for computing spatial transformations of images in registration problems.

Future work in this area could include investigating further which basis functions best preserve the covariance properties of the images during image registration while keeping in mind their computational cost. Lastly, we note that intensity based image registration is not the only application which requires continuous approximations of discretely sampled image data. Other important applications such as image segmentation using deformable models and sub-pixel edge detection and target tracking also make use of continuous approximation models based on noisy samples. Variational energy minimization methods are commonly used to compute solutions for such problems and we postulate that artifacts similar to those explained in this chapter could be encountered in these applications as well. Future work could also include demonstrating the effects of

low order interpolators in these applications.

## 4.7 Appendix A

Here we give the definition of the several interpolating basis functions used throughout this paper. The linear basis function, also known as the B-spline of order 1 or hat function, is given by:

$$h(x) = \begin{cases} 1 - |x| & \text{if } |x| \leq 1; \\ 0 & \text{if } |x| > 1. \end{cases} \tag{4.32}$$

The cubic cardinal spline function referred to above is given by:

$$h(x) = \frac{-6\alpha}{(1-\alpha^2)} \sum_{k=-\infty}^{+\infty} \alpha^{|k|} \beta^3(x-k), \tag{4.33}$$

where $\alpha = 0.2679$, and $\beta^3(x)$ is the popular B-spline of order 3:

$$\beta^3(x) = \begin{cases} \frac{2}{3} - \frac{1}{2}|x|^2(2-|x|) & , 0 \leq |x| < 1; \\ \frac{1}{6}(2-|x|)^3 & , 1 \leq |x| < 2 \ ; \\ 0 & , 2 \leq |x| \ . \end{cases} \tag{4.34}$$

Finally, the popular sinc basis function is given by:

$$h(x) = \text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}. \tag{4.35}$$

The truncated version of (4.35) is given by multiplication of (4.35) with a window function $w(x)$:

$$w(x) = \begin{cases} v(x) & \text{if } |x| < W; \\ 0 & \text{if } W \leq |x|. \end{cases} \tag{4.36}$$

In this work we use the Hann function (sometimes referred to as Hanning) defined as:

$$v(x) = 0.5 + 0.5\cos(\frac{\pi x}{W}). \tag{4.37}$$

## 4.8 Appendix B

Here we show that

$$\sum_{\mathbf{i}=-\infty}^{+\infty} [\text{sinc}(\mathbf{x} - \mathbf{i})]^2 = 1 \ \forall \ \mathbf{x}. \tag{4.38}$$

We proceed by showing the 1D version of (4.38). The $\mathbb{R}^d$ result follows naturally by taking the limits in each dimension separately. We use the following definition for the continuous Fourier transform of a continuous, integrable function $s(x)$:

$$\mathcal{F}\{s\}(u) = \int s(x)e^{\jmath 2\pi xu}dx. \tag{4.39}$$

Then we have:

$$\mathcal{F}\{\text{sinc}(x)\}(u) = \begin{cases} 1 & \text{if } |u| \leq \frac{1}{2}; \\ 0 & \text{elsewhere.} \end{cases} \tag{4.40}$$

In addition:

$$\mathcal{F}\{[\text{sinc}(x)]^2\}(u) = \begin{cases} 1 - |u| & \text{if } |u| \leq 1; \\ 0 & \text{if } |u| > 1. \end{cases} \tag{4.41}$$

The sum in (4.38) can be written as a continuous integral by:

$$\sum_{\mathbf{i}=-\infty}^{+\infty} [\text{sinc}(x-i)]^2 = \int [\text{sinc}(y)]^2 \text{comb}(x-y)dy \tag{4.42}$$

where

$$\text{comb}(x) = \sum_{i=-\infty}^{\infty} \delta(x-i) \tag{4.43}$$

and $\delta$ is the Dirac Delta Function. Using Plancherel's theorem the integral (4.42) becomes:

$$\int [\text{sinc}(y)]^2 \text{comb}(x-y)dy = \int \mathcal{F}\{[\text{sinc}(x)]^2\}(u)\text{comb}(u)e^{j2\pi xu}du. \tag{4.44}$$

Evaluating the integral above, for any given $x$, we obtain 1.

## 4.9 Appendix C

Let the probability density function (*pdf*) for the intensity values in an image be defined as a mixture of Gaussians:

$$\text{pr}_S(s) = \sum_{i=1}^{N} \alpha_i \text{pr}_{S_i}(s),\ \text{pr}_{S_i}(s) \sim N(\mu_i, \sigma^2). \tag{4.45}$$

A pictorial description of the mixture model above for a one dimensional signal is given in figure 4.23. The goal here is to investigate the entropy of the mixture model that describes a translated version of the image computed with $S^c(x+t) = \sum_{i=1}^{P} S(i)h(x+t-i)$. The variance of each component $i$ of the mixture model, in the case of an uncorrelated random process, is modified as

$$\text{pr}_{S_i^c}(s) \sim N(\mu_i, \sigma^2(t)),\ \sigma^2(t) = \sigma^2 \sum_{i=-\infty}^{\infty} \left[h(t-i)\right]^2. \tag{4.46}$$

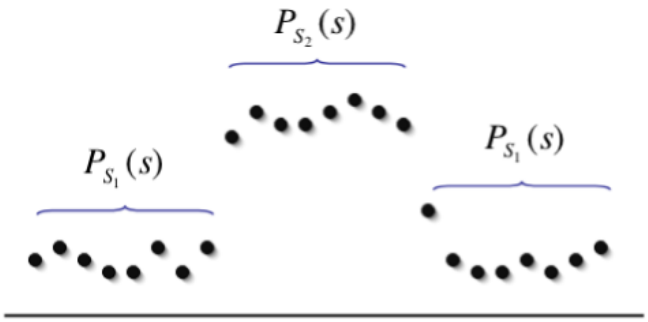


Figure 4.23: Pictorial representation of a mixture model *pdf* for a one dimensional signal. This particular example contains two classes: background and foreground.

Note that if partition of unity is satisfied $\sum_{i=-\infty}^{\infty} h(x-i) = 1, \forall x$, then $\mu_i(t) = \mu_i$. So

$$\mathrm{pr}_{S^c}(s) = \sum_{i=1}^{N} \alpha_i \mathrm{pr}_{S_i^c}(s), \;\; \mathrm{pr}_{S_i^c}(s) \sim N(\mu_i, \sigma^2(t)). \tag{4.47}$$

Note that $\sigma^2(t) \leqslant \sigma^2$, with equality when $h$ =sinc, as well as other basis functions such as Haar.

Again we wish to compare the entropy of two mixture modes: $\mathrm{pr}_S(s)$ refers to the *pdf* of the original data, and $\mathrm{pr}_{S^c}(s)$ refers to the *pdf* of the interpolated (translated) continuous signal. Then, for a fixed translation $t$ we can find $a^2$ so that:

$$\sigma^2(t) + a^2 = \sigma^2. \tag{4.48}$$

The *pdf* of the original data can be written as a function of the *pdf* of the translated signal:

$$\mathrm{pr}_S(s) = \mathrm{pr}_{S^c}(s) * N(0, a^2), \tag{4.49}$$

where $*$ refers to the one dimensional convolution operation. Let $Y_S$, $Y_{S^c}$, and $Y_a$ be random variables associated with *pdf*s $\mathrm{pr}_S(s)$, $\mathrm{pr}_{S^c}(s)$, and $N(0, a^2)$, respectively. Then:

$$Y_S = Y_{S^c} + Y_a, \tag{4.50}$$

and

$$H(Y_S) = H(Y_{S^c} + Y_a) \geqslant H(Y_{S^c}). \tag{4.51}$$

In words, the entropy of the *pdf* associated with the experiment of choosing intensity values from an image translated using a continuous image model is monotonic with respect to the variance of the signal: the more interpolation on image values, the lower the variance of the signal, the lower the entropy. Since the variance of any component in the mixture model defined in 4.45 oscillates with respect to translation parameter $t$ according to $\sigma^2 \sum_{i=-\infty}^{\infty} [h(t-i)]^2$, the entropy of a translated signal will also oscillate accordingly.

### 4.9.1 Joint entropy between two images

Mutual information-based registration of two images $S$ and $T$ also involves the estimation of their joint entropy, which in turn requires knowledge of the joint *pdf* of the images' intensity values $H(S^c, T)$,: $I(S,T) = H(T) + H(S^c) - H(S^c, T)$. A very similar argument to the one made above can be made by extending the 1D mixture model to a 2D mixture model for the joint distribution of the images:

$$\mathrm{pr}_{S,T} = \sum_{i=1}^{N} \alpha_i \mathrm{pr}_{(S,T)_i}(s,t),\ \mathrm{pr}_{(S,T)_i}(s,t) \sim N(\mathbf{u}_i, \Sigma), \tag{4.52}$$

where

$$\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \lambda^2 \end{pmatrix}. \tag{4.53}$$

Again the mixture model associated with the translated signal is:

$$\mathrm{pr}_{S^c,T} = \sum_{i=1}^{N} \alpha_i \mathrm{pr}_{(S^c,T)_i}(s,t),\ \mathrm{pr}_{(S^c,T)_i}(s,t) \sim N(\mathbf{u}_i, \Sigma^c), \tag{4.54}$$

with $\sigma^2(t) \leqslant \sigma$.

Thus we find ourselves in the same situation as before, comparing the entropy of the joint mixture model of a translated image with another image, and the entropy of the joint mixture model of the original, unprocessed, images. Now the joint *pdf* of the original data can also be written as a function of the *pdf* of the translated data:

$$\mathrm{pr}_{S,T}(s,t) = \mathrm{pr}_{S^c,T}(s,t) * N(0,a^2)(s) \tag{4.55}$$

where the convolution is only performed along the $s$ variable. By identical arguments:

$$H(S,T) \geqslant H(S^c,T). \tag{4.56}$$

This means that $H(S^c,T)$ also oscillates as the image $S^c$ is translated by $t$, just as $H(S^c)$ does.

### 4.9.2 Do oscillations in $H(S^c)$ and $-H(S^c,T)$ cancel out?

If $\mathrm{pr}_{S^c,T}(s,T) = \mathrm{pr}_{S^c}(s)\mathrm{pr}_T(t)$, for all translation values $t$, the oscillations in $H(S^c)$ and $-H(S^c,T)$ cancel out, since $I(S^c,T) = H(T) + H(S^c) - H(S^c,T)$. However, this also means that as far as the mutual information is concerned, the images do not align, since the marginal *pdf* of the images being registered are independent from one another, no matter the translation value $t$. Consider the translation interval $t \in [0,1]$ (the sub-pixel translation interval). For optimization to be successful the similarity measure $I(S^c,T)$ along this interval should be void of local optima. This requires that the derivative $\frac{d(H(S^c)-H(S^c,T))}{dt}$ must not change sign (unless it is an actual global optimum, of course). This means that

$$\frac{dH(S^c,T)}{dt} < \frac{dH(S^c)}{dt} \tag{4.57}$$

on the ascending part of the curve, for example, must hold for all $t$ in that interval. The opposite should hold on the descending part of the curve. This means that the changes in image translation value $t$ must produce sufficient negative changes in the joint entropy $H(S^c, T)$ to offset the systematic variations in both $H(S^c, T)$ and $H(S^c)$.

In situations of low signal to noise ratios $H(S^c, T)$ may not change appreciably with respect to changes in $t$ since not enough signal may be present for detecting significant statistical dependency, this causing the aforementioned 'grid' (local optima) artifacts in mutual information similarity measure computations.

In the mixture models above we used the assumption that the variance for each component of the mixture was the same. This allowed us to proceed with the arguments in a straight forward manner. However, as explained earlier in the chapter, the constant variance assumption does not hold for many important imaging modalities such as MRI. Below we extend the mixture model used above to include another component relating to the background of the image.

### 4.9.3 Mixture model for MRI

Thermal noise in MRI is commonly assumed to be Gaussian distributed [44], zero mean, variance $\sigma^2$, additive to both real and imaginary receiver channels. In magnitude reconstructed images, $A = \sqrt{A_{real}^2 + A_{imaginary}^2}$, the *pdf* of the intensity value $s$ at a particular voxel is given by a Rician distribution:

$$\mathrm{pr}(s) = \frac{s}{\sigma^2} e^{-\frac{s^2+A^2}{2\sigma^2}} \beta_0 \left( \frac{As}{\sigma^2} \right) \tag{4.58}$$

where $\beta_0$ represents the zeror order Bessel function of the first kind.

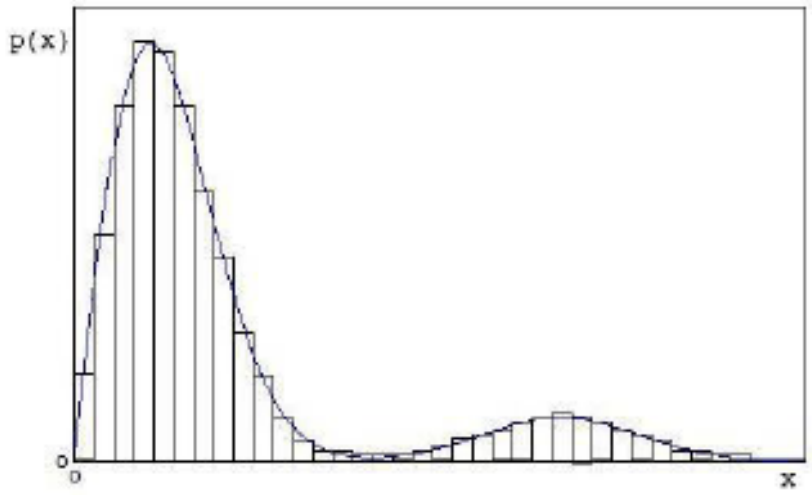


Figure 4.24: Multimodal distribution for a typical MR image.

In the low signal to noise (SNR) limit (SNR) $A/\sigma^2 \to 0$ and

$$\mathrm{pr}(s) \sim \frac{s}{\sigma^2} e^{-\frac{s^2}{2\sigma^2}}, \tag{4.59}$$

while at high SNR $A/\sigma^2 \to \infty$ and

$$\mathrm{pr}(s) \sim \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(s-A)^2}{2\sigma^2}}. \tag{4.60}$$

Note that it is commonly assumed that at SNR ratios $A/\sigma > 3$ [43] the *pdf* for any class of tissue is given by the Normal distribution specified above. Consequently, the *pdf* of intensity values in a typical MR image looks like the one displayed in Figure 4.24. The relative heights of the two most prominent peaks are given by the relative size of the foreground with respect to the background, while the distance of separation between them is given by the SNR.

Thus, a good approximation for the intensity distribution of a typical MR image, denoted here as image $S$, is the following mixure model:

$$\mathrm{pr}_S(s) = \tau \mathrm{pr}_{bkg}(s) + (1-\tau)\mathrm{pr}_{frg}(s), \tag{4.61}$$

where $\mathrm{pr}_{bkg}$ is given by the *pdf* of the background (4.59), and $\mathrm{pr}_{frg}(s)$ (the *pdf* of the foreground part) is given by the mixture model in equation (4.46). Using this model, the entropy of the *pdf* of the image is:

$$H(P_S(s)) = \int \left(\tau \mathrm{pr}_{bkg}(s) + (1-\tau)\mathrm{pr}_{frg}(s)\right) \log\left(\tau \mathrm{pr}_{bkg}(s) + (1-\tau)\mathrm{pr}_{frg}(s)\right) ds. \tag{4.62}$$

At high SNR we can approximate the above integral with:

$$H(\mathrm{pr}_S(s)) \sim \tau H(\mathrm{pr}_{bkg}(s)) + (1-\tau) H(\mathrm{pr}_{frg}(s)). \tag{4.63}$$

This means that, in MRI at high SNR, we can expect that our analysis for entropy oscillations, as done above, will follow through. That is, each term in the equation above should oscillate (it is easy to check that the entropy of the *pdf* of the background values oscillates in the same manner), and so does their sum. The same can be easily shown for the joint histogram part as well.

# Acknowledgments

We would like to thank Dr. Carlo Pierpaoli of the Section on Tissue Biophysics and Biomimetics, Laboratory of Integrative and Medical Biophysics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA for providing some of the images used in this study. Professor Berenstein's work has been partly supported by NSF grant DMS0400698.

We would also like to thank Professor Kagan of the Mathematics Department, University of Maryland, for his input in some of the ideas presented in the appendix C of this chapter.

# Chapter 5

# Comprehensive Approach for Correction of Motion and Distortion in Diffusion Weighted MRI

## 5.1 Introduction

The MRI measurement of water diffusion provides important information about compositional, structural, and organizational features of biological tissues. Most clinical MRI diffusion studies are performed by acquiring single shot echo-planar images (EPI) with diffusion sensitizing gradients of different strengths and orientations. In diffusion tensor imaging (DTI) [17], at least seven images must be acquired, one image with no diffusion sensitization, and six diffusion weighted images (DWIs) with diffusion sensitization magnetic field gradients applied in non-collinear directions, and possibly of different magnitude. In the DTI literature, the magnitude of the diffusion weighting gradient is normally denoted by a so called 'b-value' of units s/mm$^2$. Throughout this chapter we will also refer to a so called 'b-matrix' associated with a particular diffusion sensitizing magnetic field gradient, which is simply proportional to the outer product of the magnetic field gradients. For more details, please refer to [17]. DTI analysis, as well as other approaches to extract diffusion information from MRI data

[113, 119, 57, 36], require the different DWIs to be spatially co-registered.

Unfortunately, eddy-current-induced image distortions and patient motion during prolonged acquisitions cause misalignment of the diffusion weighted images (DWIs). Eddy currents are significantly reduced, but not eliminated, by actively shielded gradients [2] and pre-emphasis correction schemes [1] in modern magnets. Methods proposed to reduce residual eddy-current-induced distortions in DWIs are either field map or image based.

In a field-map-based correction scheme, such as that presented by Jezzard et al [58], one measures the magnetic field produced by the eddy currents and then corrects the distortion using the field map and theoretical models of how field inhomogeneities distort the images. The major obstacle to implementation is the difficulty of rapidly acquiring reliable field maps.

In an image-based registration scheme, one uses a cost function $Q$ to measure how well the images are spatially aligned. First, a target image is chosen as a reference for all other images in the data set (source images). Because it is usually less distorted and has a higher signal-to-noise ratio than the heavily diffusion-weighted images, the image acquired with no diffusion sensitization (T2WI), is usually used as the target image for registering DWIs. Next, using a spatial transformation model, one aligns all other images to the target image by optimizing a cost function. Image-based registration schemes differ from each other in: 1) the definition of $Q$, 2) the types of transformations applied to the image in searching for the maximum of $Q$, and 3) the numerical optimization method used in searching for the maximum of $Q$. We chose to work with an image-based registration scheme because a field map, which is usually not available for DWIs acquired on a conventional clinical scanner, is not required. In addition, an image-based

scheme allows for correction for misregistration produced by subject motion.

Haselgrove [47] proposed the first image-based registration method to correct for eddy- current-induced distortions. He used the undistorted T2WI as a target image for the registration of the DWIs. $Q$ was based on the cross correlations between source image and target image. Unfortunately, cross correlation performs poorly as a measure of alignment when the contrast of source and target images differs significantly. Experiments by Bastin et al. [20] indicate that this approach does not perform well in registering T2WIs to DWIs acquired with b-values higher than 300 s/mm$^2$.

Cost functions based on mutual information are more robust than those based on correlation for registering images with significantly different contrast. A mutual information based method was presented by Horsfield [55]. He proposed to measure the effect of eddy currents by registering DWIs acquired on a special phantom having low diffusivity and non-uniform relaxation properties. The correction parameters computed from the phantom can in turn be used to correct anatomical DWIs acquired in the same scanner with the same sequence provided that 1) the subjects position in the scanner is sufficiently well known, and 2) gradient performance is stable over time.

Andersson and Skare recently proposed a scheme that uses the goodness-of-fit of the apparent diffusion tensor (D) calculation for guiding the image registration process [8]. They define $Q$ using the goodness-of-fit $\chi^2$ of the DWI signal to the b-matrix [17]. This interesting approach has the advantage of correcting for both eddy-current-induced distortions and subject motion. As the authors point out, however, this method cannot be used to register DWIs to T2WIs when a single level of diffusion weighting is sampled because, in this case, the T2WIs do not

contribute to $\chi^2$.

Surprisingly, until recently [8] misregistration from patient motion has been neglected as a potential source of artifacts in diffusion MRI studies, although it has been long recognized as such in functional MRI studies.

The goal of this work was to design a robust image registration approach that would correct the spatial misregistration of DWI volumes originating from both subject motion and eddy-current-induced distortions. Our post-processing method can also be used to position the dataset in a standardized orientation. The b-matrix and the signal magnitude of each DWI volume are recalculated to take into account the effects of the spatial transformation applied. The method requires only one image interpolation step, thus avoiding unnecessary blurring of the images, and without requiring additional measurements on phantoms or additional scans to map the magnetic field produced by the eddy currents. Note that most of the work presented in this chapter was previously published, with some modifications, in Rohde *et al.* [91].

## 5.2 Materials and Methods

In this section, we first describe the pulse sequence and the acquisition parameters we used. We then present a mathematical formulation of the registration problem. We describe the spatial transformation $f$, the cost function $Q$, the numerical method for finding the model parameters that maximize $Q$, and the post-registration processing necessary to correct the image brightness and to rotate the b-matrices. Finally, we describe the experimental design and data analysis approach for the tests we performed to validate our results.

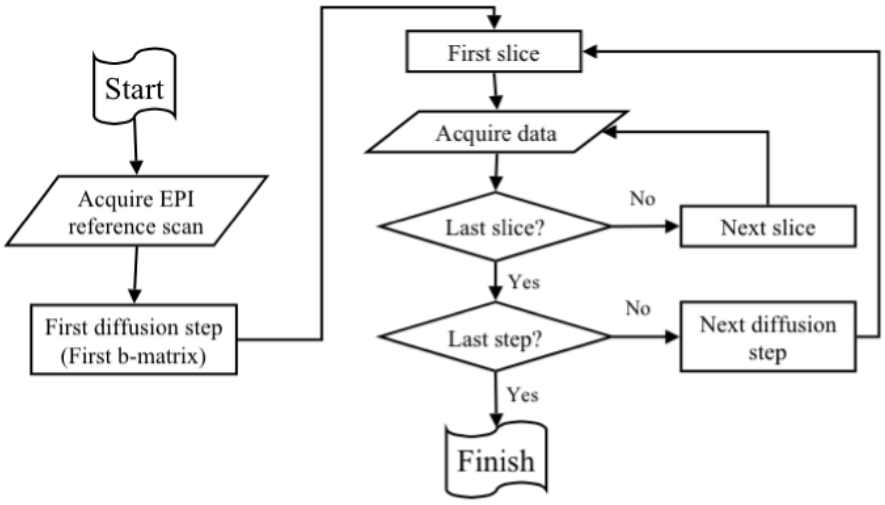


Figure 5.1: Block diagram description of the volume acquisition procedure.

### 5.2.1 Pulse sequence and MRI parameters

We designed our correction scheme to reduce artifacts in data sets acquired with a standard single-shot multi-slice spin-echo EPI sequence (i.e.: fat suppression pulse, 90 degree pulse, first diffusion gradient, 180 degree pulse, second diffusion gradient, EPI readout). The acquisition order is shown in Figure 5.1. We begin with a dummy scan, acquired with no radio frequency (RF) excitation, which would permit us to correct for the direct current offset of the RF amplifier. Next we acquire an EPI reference scan for each slice. Finally we collect the image data looping through all slices and the different diffusion weightings. The slice loop is the innermost loop, so we acquire all slices composing a volume before proceeding to the next diffusion weighting.

Scans were performed on a 1.5 T GE Signa system equipped with a whole-body gradient coil able to produce gradient pulses up to 50 mT/m (GE Medical Systems, Milwaukee, WI). We scanned a cylindrical silicone oil phantom and the brains of healthy subjects. The pre-emphasis correction for eddy currents

provided by the system manufacturer was disabled in some scans to obtain images with severe geometric distortions. Some imaging parameters, such as resolution, repetition time, total number of images, etc., varied slightly for the different tests we performed. For brevity we report the imaging parameters of the brain study shown in Figures 5.7, 5.8, 5.9, 5.10. The imaged volume was composed of 80 contiguous slices with 2 mm slice thickness and 2 mm in-plane resolution. The echo-time was 82.7 ms, the read-out time 50 ms, and the repetition time was greater than 10 s with cardiac gating (4 acquisitions per heart beat starting with a 150 ms delay after the rise of the sphygmic wave as measured with a peripheral pulse oxymeter). The gradient strength was 49 mT/m, yielding a b value (i.e., trace of the b-matrix) of 1,120 s/mm$^2$. A total of 56 3D images were acquired by repeating 8 times a diffusion sampling scheme described previously (23) which includes one volume with no diffusion weighting followed by the same volume six times, acquired with diffusion gradients applied in different directions. The total imaging time was approximately 20 minutes. Replicate volumes were acquired for signal to noise considerations in order to improve the quality of the estimated diffusion tensor parameters.

### 5.2.2 Formulation of the spatial transformation model

A data set from a diffusion MR study consists of multiple volumes acquired with different strengths and orientation of the diffusion sensitizing gradients. To register a set of images or 3D volumes, we first define a target coordinate system $\mathbf{x}$. We then define a source coordinate system $\mathbf{x}_\alpha$ for each volume $\alpha$. The registration problem consists of finding the coordinate transformation $f_\alpha(\mathbf{x})$ for each volume $\alpha$ that transforms the target coordinates $\mathbf{x}$ into the source coordinates $\mathbf{x}_\alpha$. We

describe the transformations with a set of parameters $\mathbf{p}_\alpha = \{p_\alpha^1, p_\alpha^2, \cdots, p_\alpha^m\}$, where $m$ is the total number of parameters:

$$\mathbf{x}_\alpha = \mathbf{f}_\alpha(\mathbf{x}) = \mathbf{f}_\alpha(\mathbf{x}; \mathbf{p}_\alpha). \tag{5.1}$$

We then define a cost function $Q$ that depends upon the image values and the parameters. We then find $\mathbf{f}$ by finding $\mathbf{p}_\alpha$ for which $Q$ is an extremum.

The first T2-weighted volume acquired is extracted from the DWI dataset and rigidly registered to a standard template. Subsequent registrations of each DWI to the normalized T2-weighted target volume will then cause then entire DWI dataset to be positioned in a standardized orientation. All human brain images displayed in this paper were spatially normalized using this method. Although this volume is free from eddy-current-induced distortions, it contains geometrical distortions due to $\mathbf{B}_0$ inhomogeneity caused by the magnetic susceptibility of the object. These distortions are generally accompanied by localized changes in the brightness of the image. As opposed to eddy currents caused by the diffusion weighting magnetic gradients, which can vary with the different gradients being applied, the amount of $\mathbf{B}_0$ inhomogeneity is likely to remain constant, since it does not vary with any acquisition parameters. Since distortions due to $\mathbf{B}_0$ inhomogeneity are the same in both T2WIs and DWIs, they do not cause misregistration artifacts in the computed diffusion parameters. However, if large corrections of the shape of the object (in particular magnification) are required to correct the DWIs for eddy-current distortion, one should also correct their brightness (signal magnitude) appropriately to account for changes in size of the object. Neglecting to correct the brightness of the DWIs following distortion correction will result in errors in the computed diffusion parameters.

Each volume $\alpha$ consists of a 3D array of pixel values, $I_\alpha[p,q,r]$ where $p$, $q$, and $r$ are the array indices. We define the image coordinates for volume $\alpha$ as

$$\mathbf{x}_\alpha = (x_{1\alpha}, x_{2\alpha}, x_{3\alpha}) = (s_1 p, s_2 q, s_3 r) \tag{5.2}$$

where $s_1$ and $s_2$ are the nominal in-plane pixel sizes and $s_3$ is the slice separation. Each volume $I_\alpha[p,q,r]$ is only defined for integer values of the indices $p$, $q$, and $r$. We use tri-linear interpolation [86] to define a continuous function $I_\alpha(\mathbf{x}_\alpha)$ in terms of the measured values.

Call the coordinate system of the target image $\mathbf{x}$. For each volume $\alpha$ we seek the properly registered and brightness-corrected image $I^r_\alpha$. As shown in equation [8], $I^r_\alpha$ is related to the measured image $I_\alpha$ by the equation

$$I^r_\alpha(\mathbf{x}) = I_\alpha\left(\mathbf{f}_\alpha(\mathbf{x})\right)\Gamma\left(\mathbf{f}_\alpha(\mathbf{x})\right), \tag{5.3}$$

where $\mathbf{f}_\alpha$ is a coordinate transformation and $\Gamma_\alpha(\mathbf{x})$ is a brighness correction function. We propose a parameterized form for $\mathbf{f}_\alpha$ that can correct for patient motion and eddy-current- induced distortion. We find the best values of the parameters by maximizing the cost function $Q$ as described below. Since the brightness correction is a function of the spatial transformation necessary to correct for eddy-current-induced distortions, we can express the brightness correction $\Gamma_\alpha(\mathbf{x})$ in terms of parameters we used to model the eddy current field (See: Effects of eddy currents section below).

We decompose the spatial transformation $\mathbf{f}$ into two steps. The first step describes the change in location and orientation of the object between acquisition times of the two volumes and the second step describes the distortion introduced by the acquisition process.

We model the patients brain as a rigid body and describe its displacement and change in orientation by the equation

$$\mathbf{y}_\alpha = \mathbf{R}\mathbf{x} + \mathbf{a} \tag{5.4}$$

where $\mathbf{R}$ is a rotation matrix, $\mathbf{a}$ is a displacement vector, and $\mathbf{y}_\alpha$ is an intermediate set of coordinates that describes the orientation of the patients head at the time of the acquisition of volume $\alpha$. Equation (5.4) contains six parameters, the three components of the displacement vector $\mathbf{a}$ and three parameters, typically Euler angles $\varphi$, $\theta$, and $\psi$ that define the rotation $\mathbf{R}$ [40]. No changes in the image brightness are associated with this transformation.

Subject motion between the start of the excitation pulse and the end of the data acquisition can potentially affect the data in a significant way. For a single shot EPI acquisition, this period lasts about 1.5 TE, or about 120ms in our case. Head motion during this period has two effects on the data. 1) The component of velocity parallel to the diffusion sensitizing gradients causes a phase shift in the reconstructed image. This phase shift does not cause any artifact in EPI single shot magnitude images. 2) Motion during the readout (the integration process during image formation, such as in equation 5.23, for example) of can cause blurring. An image registration algorithm is clearly not appropriate for correcting image blurring.

### 5.2.3 Effects of eddy currents

The image distortion due to eddy currents depends on the time and space dependence of $b_{\text{eddy}}(\mathbf{y}_\alpha, t)$, the component parallel to $\mathbf{B}_0$ of the magnetic field generated by the eddy currents. The shifts in the readout, phase encode, and

slice select directions are

$$\delta_{\text{read}} = \frac{b_{\text{eddy}}}{g_{\text{readout}} s_1} \cong \frac{b_{\text{eddy}}}{7\text{gauss}} \tag{5.5}$$

$$\delta_{\text{phase}} = \frac{b_{\text{eddy}} \tau_{\text{echo}}}{G_{\text{blip}} s_2} \cong \frac{b_{\text{eddy}}}{0.7\text{gauss}} \tag{5.6}$$

$$\delta_{\text{slice}} = \frac{b_{\text{eddy}}}{g_{\text{slice}} s_3} \cong \frac{b_{\text{eddy}}}{2.5\text{gauss}} \tag{5.7}$$

where $\delta_{\text{read}}$ and $\delta_{\text{phase}}$ are in pixels, $\delta_{\text{slice}}$ is in multiples of the slice thickness, $G_{\text{blip}}$ is the area of the phase encode blips, $\tau_{\text{echo}}$ is the time between consecutive echoes in the readout echo train, $s_1$, $s_2$, and $s_3$ are defined in equation (5.2), and the numerical values are typical of the scans we perform. We can safely ignore the shifts in the readout and slice select directions, as they are almost two orders of magnitude smaller than the shifts in the phase encode direction. If $b_{\text{eddy}}(\mathbf{y}_\alpha, t)$ were to change appreciably during the readout phase of image acquisition, the image would suffer blurring that cannot be corrected by a simple coordinate transformation.

As shown in appendix A, if we assume that $b_{\text{eddy}}(\mathbf{y}_\alpha, t)$ is approximately constant during the readout and that eddy current fields from the acquisition of one slice do not interfere with the next, the pixel values for volume $\alpha$ are

$$I_\alpha[m,l,q] = \tilde{W}_\alpha \left( m s_1, l s_2 - \beta b_{\text{eddy}}(m s_1, l s_2, q s_3), q s_3 \right) \left( 1 + \frac{\partial(\beta b_{\text{eddy}})}{\partial y_2} \right) \tag{5.8}$$

where $\tilde{W}_\alpha$ is the undistorted image of volume $\alpha$ smoothed by the point spread function of the imaging system, and $\beta$ is given in equation 5.25. Comparing equations (5.8) and (5.3) we conclude that the transformation and brightness

correction associated with eddy currents are

$$\mathbf{x}_\alpha = \mathbf{y}_\alpha - \beta b_{\text{eddy}} \mathbf{e}_2 \tag{5.9}$$

with $\mathbf{e}_2 = (0, 1, 0)$ representing the phase encode direction and

$$\Gamma(\mathbf{x}_\alpha) = \left( 1 + \frac{\partial(\beta b_{\text{eddy}})}{\partial y_2} \right)_{\mathbf{x}_\alpha = \mathbf{y}_\alpha} \tag{5.10}$$

We now have to model $b_{\text{eddy}}$. Since we can neglect the fields due to eddy currents induced in the patients head, $b_{\text{eddy}}(\mathbf{y}_\alpha, t)$ in the imaging volume obeys Laplaces equation [102]

$$\nabla^2 b_{\text{eddy}} = 0 \tag{5.11}$$

Expanding the solution of Laplaces equation in Cartesian coordinates up to second order, we approximate $b_{\text{eddy}}$ by

$$w(x) = \begin{array}{l} c_0 + c_1 y_{\alpha 1} + c_2 y_{\alpha 2} + c_3 y_{\alpha 3} + c_4 y_{\alpha 1} y_{\alpha 2} + c_5 y_{\alpha 1} y_{\alpha 3} + \\ c_6 y_{\alpha 2} y_{\alpha 3} + c_7 (y_{\alpha 1}^2 - y_{\alpha 2}^2) + c_8 (2 y_{\alpha 3}^2 - y_{\alpha 1}^2 - y_{\alpha 2}^2) \end{array} \tag{5.12}$$

where $c_0$-$c_8$ are parameters to be determined from the optimization procedure. We can use equation (5.12) to write the brightness correction function in terms of the fit parameters:

$$\Gamma_\alpha = 1 + c_2 + c_4 y_{\alpha 1} + c_6 y_{\alpha 3} + 2(c_7 + c_8) y_{\alpha 2} \tag{5.13}$$

Equations (5.4),(5.9), and (5.12) define the transformation $\mathbf{f}_\alpha$ from target coordinates $\mathbf{x}$ to the source coordinates $\mathbf{x}_\alpha$. This transformation is not unique, however, because the $c_0$ and $t_2$ are not independent. We therefore set $c_0 = 0$ without loss of generality, and are left with 14 parameters: eight ($c_1$-$c_8$) that describe the eddy current distortion and six ($a_1, a_2, a_3, \varphi, \theta,$ and $\psi$) that describe the rigid body displacement of the object.

As mentioned previously, very short time constant eddy currents (i.e., eddy currents with significant variation during the read-out period) do not produce image distortion but rather image blurring that cannot be corrected by image registration. Our approach will correct for the effect of eddy currents that have relatively long time constants with negligible decay during the read-out period. However, our 3D correction model assumes that the eddy-current field is the same for all the slices in a particular volume. This assumption requires either that the eddy currents from the previous excitation have died away during the time interval between consecutive excitations or that a steady state is reached. In the latter case, the first few slices acquired in each volume after the gradient direction has changed will be collected before the steady state is reached and will have an amount of distortion inconsistent with that accounted for by our model. We performed a set of 2D registration experiments on a silicone oil phantom in order to investigate this effect on our magnet.

### 5.2.4 Cost function

In image registration problems, the goal of the cost function $Q$ is to measure how well two images are aligned. It is common to assume that the images are optimally aligned when the statistical dependence between their intensity values is highest. As mentioned earlier, the correlation coefficient is a poor measure of image alignment when the intensities in the images are not linearly related (see results section). A more robust way of measuring spatial alignment in medical images is to use the mutual information, a special case of the Kullback-Leibler measure [62], between the intensity values of the images to be registered [118, 56, 70]. The mutual information similarity measure has been shown to be significantly

more robust than the correlation coefficient and some other measures of similarity, when registering images of different modalities [105]. In practice, we elected to use the normalized mutual information [104] in our registration program as it has been shown to avoid any dependency on the amount of image overlap.

Consider two volumes $S$ and $T$. The normalized mutual information is defined in terms of three quantities: the normalized histogram $p_S(n)$ of pixel values in volume $S$, the normalized histogram $p_T(m)$ of pixel values of image $T$, and the normalized joint histogram $p_{ST}(nm)$ of pixel values in volume $T$ and the corresponding pixel in image $S$. Let $v_{min}$ and $v_{max}$ be the minimum and maximum pixel values in volume $S$ and $w_{min}$ and $w_{max}$ be the minimum and maximum pixel values in image $T$. We divide the range of pixel values of volume $S$ into $N$ bins of equal width and the range of pixel values in image $T$ into $M$ bins of equal width. The limits of bin $l$ for volume $S$ are $v_l$ and $v_{l+1}$ given by

$$v_l = v_{min} + l\frac{v_{max} - v_{min}}{N} \tag{5.14}$$

and the limits of bin $m$ for image $T$ are $w_m$ and $w_{m+1}$ given by

$$w_m = w_{min} + m\frac{w_{max} - w_{min}}{M} \tag{5.15}$$

Let $n_{Sl}$ be the number of pixels in volume $S$ with value $h$ in the range $v_l \leqslant h < v_{l+1}$, $n_{Tm}$ the number of pixels in image $T$ with value $d$ in the range $w_l \leqslant d < w_{l+1}$, $n_{SlTm}$ the number of voxels for which the value $h$ in volume $S$ lies in the range $v_l \leqslant h < v_{l+1}$ and the value of the corresponding voxel in image $T$ lies in the range $w_l \leqslant d < w_{l+1}$, and let $n_{tot}$ be the total number of voxels in volume $S$ (= the total number of voxels in image $T$). The normalized histogram of volume $S$ is then

$$p_S(n) = \frac{n_{Sl}}{n_{tot}} \tag{5.16}$$

the normalized histogram of image $T$ is

$$p_T(m) = \frac{m_{Tm}}{n_{tot}} \tag{5.17}$$

and the normalized joint histogram of volumes $T$ and $S$ is

$$p_{ST}(l,m) = \frac{n_{SlTm}}{n_{tot}}. \tag{5.18}$$

Our cost function Q is defined in terms of the above histograms by the formula

$$Q(S,T) = \frac{\sum_l p_S(l)\ln(p_S(l)) + \sum_m p_T(m)\ln(p_T(m))}{\sum_m \sum_l p_{ST}(l,m)\ln(p_{ST}(l,m))}. \tag{5.19}$$

Lastly, note that often the images reconstructed from an MRI experiment contain intensity spikes, due to the possibility of artifacts which can arise during many stages of the image acquisition and reconstruction process. That is why, in practice, we do not choose $v_{max}$ and $w_{max}$ to be the maximum of images $T$ and $S$, respectively. Instead, we sort the pixels of $T$ by increasing intensity, and choose $w_{max}$ to be some $i^{\text{th}}$ (with $i$ some low integer) value from the last element in the sorted list. The same operation is performed to compute $v_{max}$.

### 5.2.5 Optimization

To register a source volume $\alpha$ to the target volume $T$ we have to find the values of the 14 parameters $\mathbf{p}_\alpha = (c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, a_1, a_2, a_3, \varphi, \theta, \text{ and } \psi)$ for which in equation (5.19) is a maximum when we use the target volume for volume $T$ and equations (5.3), (5.4), (5.9), and (5.12) to compute the volume $S$ from the measured volume $\alpha$. To speed up the registration, we do not apply the brightness correction $\Gamma_\alpha$ during the optimization process. We use a simple gradient ascent optimization procedure coupled with a golden section line optimization method

to maximize $Q$ [86]. That is, the registration parameters are searched iteratively according to the equation

$$\mathbf{p}_\alpha^{k+1} = \mathbf{p}_\alpha^k + \zeta \nabla_{\mathbf{p}_\alpha} Q(\mathbf{p}_\alpha^k), \tag{5.20}$$

where $\zeta$ is determined via the golden section line optimization method (note the change in notation $Q(S(\mathbf{f}(\mathbf{x}; \mathbf{p}_\alpha)), T(\mathbf{x})))$.

We evaluate the gradient of the cost function with respect to registration parameters $\mathbf{p}_\alpha$ numerically using finite differences, and several step sizes are used throughout the registration procedure. Note that because different sets of parameters in $\mathbf{p}_\alpha$ have different units, thus affecting the value of $Q(\mathbf{p}_\alpha)$ differently, we use equation (5.20) to determine different sets of parameters in $\mathbf{p}_\alpha$ separately. That is, equation (5.20) is used to update the translation parameters first, then rotation parameters, followed by the linear deformation parameters, with the quadratic deformation parameters last. The loop is repeated until the improvements in the cost function value fall bellow a chosen tolerance (typically $\epsilon = 0.0001$). The algorithm also works in multiple resolutions, in a coarse to fine fashion, in order to avoid local optima and decrease computation time. A flowchart describing the optimization loop is given in Figure 5.2.

### 5.2.6 Post-registration processing

#### Intensity correction

After registering the volume to the target we apply the brightness correction $\Gamma_\alpha$. An example demonstrating the benefit of the brightness correction is shown in the results section.

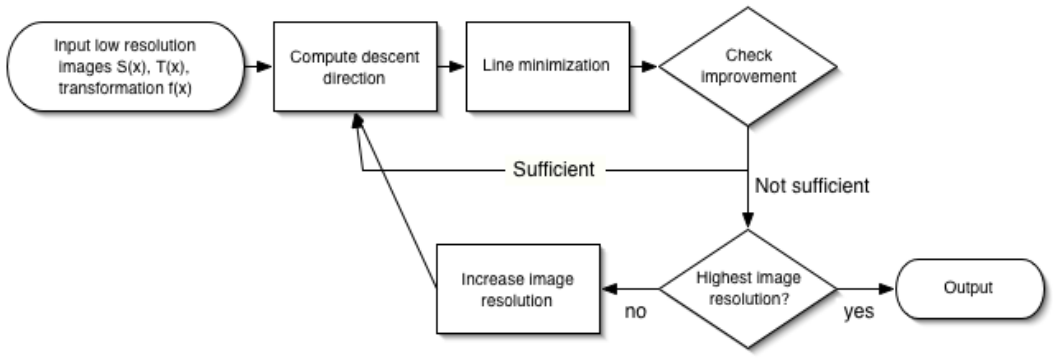


Figure 5.2: Flowchart description of source code implementation of optimization loop.

### B-matrix reorientation

Each DWI volume is accompanied by a b-matrix that describes its diffusion weighting. In principle, both patient motion and eddy currents can cause errors in calculation of parameters that describe diffusion in each voxel of a DWI dataset. We use dimensional analysis to estimate the change in the b-matrix due to the eddy currents and to demonstrate that it is very small in normal conditions. The elements of the b-matrix scale as

$$b \propto \gamma^2 g_{diff}^2 \tau^3 \tag{5.21}$$

where $\gamma^2$ is the gyro-magnetic ratio, $g_{diff}$ is the strength of the diffusion gradients, and $\tau$ is the characteristic time for the experiments. The presence of an eddy current $b_{eddy}$ would change the b-value to $\gamma^2 \left| g_{diff} + \frac{b_{eddy}}{L} \right| \tau^3$, resulting in a fractional change in $b$ of $\frac{2b_{eddy}}{g_{diff}L}$, where $L$ is a characteristic length scale over which $b_{eddy}$ varies. For the MRI scanner used throughout this chapter, with pixel shifts of the order of 1-2 pixels, over a 22cm field of view, $b_{eddy} \approx 0.1$ gauss, $L \approx 20$ cm, and $g_{diff} \approx 4$ gauss/cm, resulting in a fractional change in the b-value of about

0.25 percent, which we can safely neglect.

We do correct for patient motion, however. The b-matrix is calculated with respect to the $\mathbf{y}_\alpha$-coordinates, which are fixed with respect to the scanner. We have to rotate the b-matrix to target coordinates $\mathbf{x}$ using the results of the fitting procedure:

$$\mathbf{b}' = \mathbf{R}^{-1}\mathbf{b}_\alpha\mathbf{R} \tag{5.22}$$

where $\mathbf{b}'$ is the rotated b-matrix in target coordinates, $\mathbf{b}_\alpha$ is the computed b-matrix, and $\mathbf{R}$ is the rotation matrix defined by the Euler angles $\varphi$, $\theta$, and $\psi$ obtained from the fitting procedure.

### 5.2.7 Validation methods

A common problem encountered in validating results from image registration algorithms, particularly nonrigid ones, is the lack of a gold standard. Therefore we are limited to using indirect measures to establish the reliability of our spatial normalization approach. In this section we describe several approaches we used to establish the reliability of the results produced by our registration algorithm.

#### Visual assessment of DWIs and computed diffusion tensor images

As an initial check, we confirm through visual inspection that DWIs and their corresponding T2-weighted images are well aligned after registration. Then we check for artifacts in the maps of the computed tensor parameters that could originate from image misalignment in DWI datasets used to generate them. Often such artifacts are large enough to be detected at visual inspection of the data. An assumption often used [58, 47, 55] is that high anisotropy index values around the

edges of the brain are associated with image misalignment because anisotropy in those regions is inconsistent with known anatomy. Following the same approach, we use anisotropy index images, as well as images of the off-diagonal tensor elements, to assess the amount of anisotropic diffusion at the periphery of the brain.

### Improved fitting to the tensor model

Andersson and Skare [8] proposed an image registration approach that uses the goodness-of-fit of the apparent diffusion tensor (**D**) calculation for guiding the image registration process. Here we use their approach to test the results of our registration algorithm. This scheme relies on the assumption that the DWI data is well described by the tensor model of Basser *et al.* [17]. At the b-values we used, this assumption is probably satisfied, although it may not hold true when very large b-values are used. Moreover, as mentioned in the introduction, this approach can only be used to test the registration of different DWI volumes; it does not provide information about the degree of registration of DWI volumes to T2WI volumes.

### PCA analysis of the data

Several factors contribute to the signal amplitude in each particular voxel of the volumes a diffusion MRI dataset. These factors include: relaxation and diffusion properties of the tissue, as well as noise. Misregistration artifacts caused by motion and image distortion will affect the signal amplitude by changing the tissue that is imaged at a particular location during different acquisitions. The $n$ volumes of a diffusion MRI dataset can be viewed as a random vector

$\mathbf{X}$ of $n$ components and $m$ elements, where $m$ is the number of voxels in the imaged volume. We use the principal component analysis (PCA) described in [7] to describe the DWI dataset in terms of a new set of $n$ uncorrelated volumes (principal components). These principal components are computed as a linear combination of the $n$ components of $\mathbf{X}$ that have special properties in terms of variances of their $m$ elements, and they are ordered by decreasing variance. For example, the first principal component is the normalized linear combination (the sum of squares of the coefficients being one) with maximum variance. The second component is the normalized linear combination that is uncorrelated with the first component and whose variance is maximal, and so on. It turns out that the coefficients of such linear combinations are given by the characteristic vectors of the covariance matrix of $\mathbf{X}$ [7]. As shown in the results section, principal components analysis is a powerful tool that can be used to detect interesting features present in a DWI sequence, including artifacts such as image distortion and patient motion.

## 5.3 Results

### 5.3.1 Cost function

Figures 5.3 and 5.4 show that correlation type similarity measures are not appropriate for registering T2WIs and DWIs. Correlation type similarity measures require signals in the target and source image to be linearly related. Figure 5.3, which displays the joint histogram of a DWI and a T2-weighted image, show that the intensity values of the images are not linearly related. This happens because cerebrospinal fluid (CSF) has low intensity in the DWIs, but not in the T2WIs.

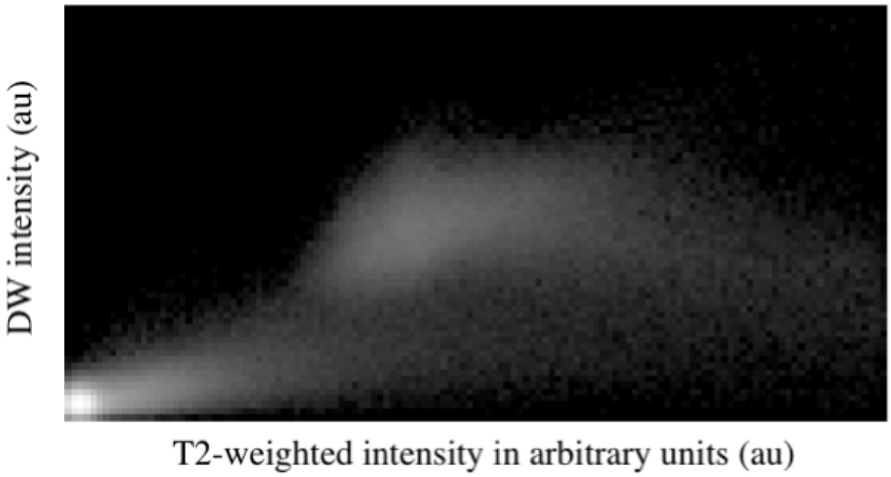


Figure 5.3: Joint histogram showing the relationship between T2 and DWI intensities. The highly nonlinear nature of this relationship demonstrates why correlation type similarity measures are not appropriate for this application. Intensities are in arbitrary units.

Consequently, background values in the DWIs map to both background and CSF values in the T2WI. In addition, DWIs contain signal affected by anisotropic diffusion, while the T2WIs do not. As a consequence, white matter has a relatively narrow range of values in the T2WIs but a large range of values in the DWIs.

Figure 5.4 shows representative results for registration of a DWI to a corresponding T2WI using the correlation coefficient as well as the normalized mutual information as the similarity measure. The first (middle image) result was obtained using the normalized mutual information cost function. The image on the right was obtained using the correlation coefficient as a similarity measure. As evident from the picture, the result obtained using the mutual information cost function is superior to the result obtained using the correlation coefficient. We note that the original DWI images were acquired with gradient pre-emphasis turned off. Thus the original images contained significant geometric distortions.

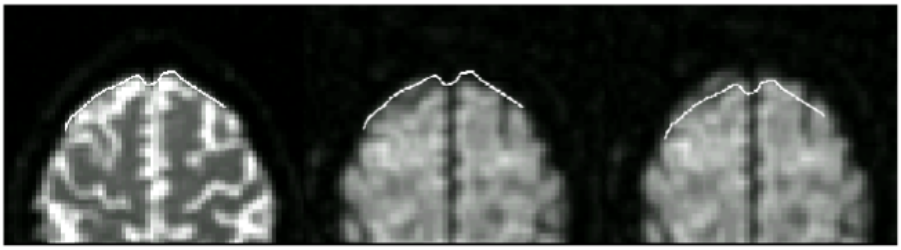

Figure 5.4: Example image registrations performed using mutual information and correlation coefficient similarity measures. Left image: the reference T2 image. Middle image: DW image registered to the reference image using normalized mutual information. Right image: DW image registered to the reference using the correlation coefficient similarity measure.

For brevity, the original distorted images are not shown here. Close visual inspection reveals that the edges of the brain shown in the images are well matched only when the normalized mutual information cost function is used. The use of the correlation coefficient causes the gray matter in the DWI to be matched to the CSF in the T2WI. Visual inspections such as these were performed as an initial accuracy check of our correction approach and revealed that the image distortion model described above seems appropriate for correcting relatively large distortions in images acquired with b-values of about 1100 $mm^2/s$, which are typically used in clinical studies.

### 5.3.2 Adequacy of the eddy current-induced distortion model

We tested whether the terms included in equation (5.12) are sufficient to approximate the distortions encountered in the image volume. We measured the distortion produced by x, y, and z oriented diffusion gradients in each slice of

the phantom using a 2D affine registration algorithm. The resulting correction coefficient for translation, magnification, and shear are plotted against slice position in Figure 5.5. The translation coefficient has dimensions of length and it is measured in pixels; the magnification coefficient is unitless, representing the ratio between the size of the original and the corrected image; and the shear coefficient is expressed in pixels per column. Slice position zero denotes the isocenter of the magnet and each slice is indexed in terms of its slice position, in millimeters. The gradient pre-emphasis scheme provided by the magnet manufacturer was turned off during the acquisition of these images and consequently large corrections are necessary to register the DWIs to the undistorted T2WI. With the exception of the first few slices, a quadratic function describes well the required correction as a function of slice position for all gradients, indicating that the terms included in equation (5.12) are sufficient to approximate the correction for distortions encountered in the image volume. Values of the correction coefficients for the first few slices differ significantly from those of the remaining slices. This suggests that the eddy-current field is not constant during the acquisition of the first few slices of a new DWI volume. In a separate experiment we performed a single slice measurement on a phantom in order to characterize the time course of the achievement of the steady state of the eddy current field for the x, y, and z gradient. This experiment showed that, in our magnet for all gradients, reaching a steady state in the eddy-current field requires about 500 to 750 ms when images are acquired with a TR of 250 ms. This result is in line with the findings of the multi slice experiment described in Figure 5.5. We add that identical experiments were performed using an image acquisition sequence in which the gradient pre-emphasis scheme was turned on. The results for these experiments

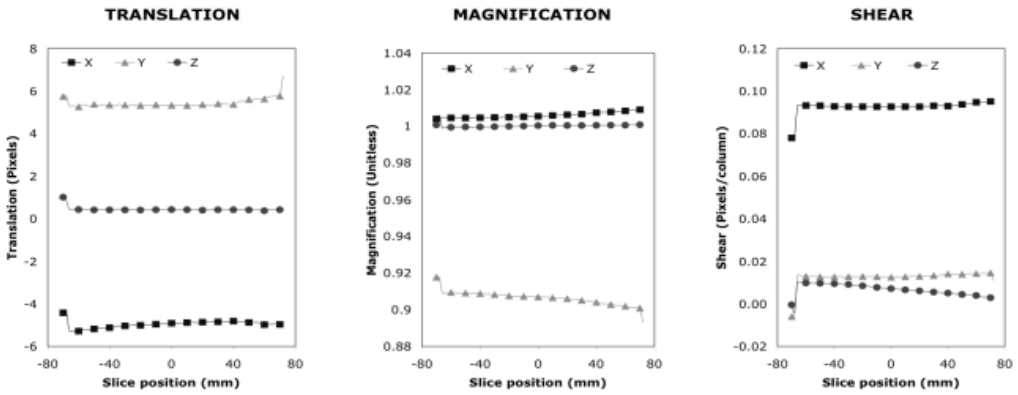


Figure 5.5: Study of image distortions using phantoms. Registration of DWIs and a target T2 image was performed slice-by-slice using an affine transformation. The registration parameters (translation, shear, and scaling) for the DWIs acquired with sensitizing gradients in the x, y, and z directions are shown above.

were qualitatively equal to the ones shown, though the distortions measured were significantly smaller. For brevity, we do not report them.

### 5.3.3 Intensity correction

The effect of omitting the intensity correction step when significant distortions are present is shown in Figure 5.6. Here a set of DWIs with severe distortion was acquired by turning off the gradient pre-emphasis eddy current compensation. These DWIs were co-registered and aligned to the T2WI and D was computed from two sets of images, one with brightness correction and one without. Figure 5.6 shows the fiber orientation color maps [81] computed from diffusion tensors obtained from these two sets of images. In the color map computed from images that had no brightness correction (left image) a red background in isotropic re-

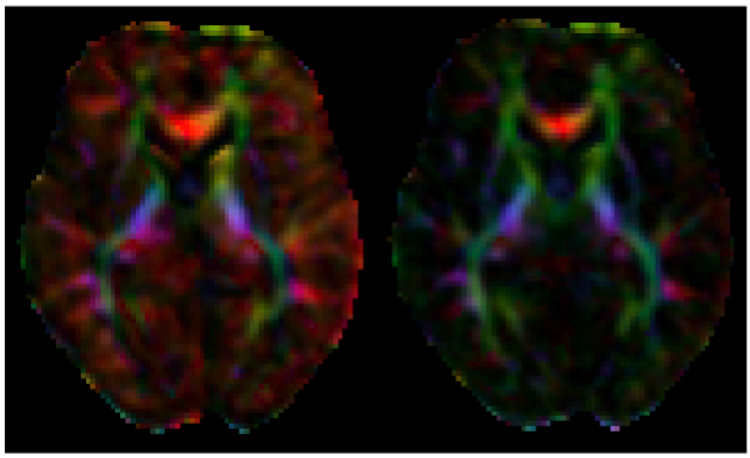

Figure 5.6: Demonstration of artifacts introduced when the intensity values of the DWI are not recomputed after warping. Left image: color representation of tensor maps computed from DWI whose intensity values were not recomputed after registration. Right image: the same, but with DWI intensity values properly recomputed. The background bias (red in color image) seen on the left image seems to indicate preferential diffusion in the x-direction. The background artifacts disappear when appropriate correction is used.

gions is evident, indicating anisotropic diffusion in the left-right orientation. This artifact is completely removed in images whose signal magnitude was appropriately corrected for the amount of non-rigid body distortion applied during image warping.

### 5.3.4 Subject motion

In order to test the ability of our algorithm to correct for subject motion, we acquired a dataset in which the subject was asked to move his head deliberately three times: about 10 degrees rotation to the left, about 10 degree rotation to the right, and an extension of about 10 degrees. The resulting DWI images in

this dataset were significantly misaligned due to patient motion. Gradient pre-emphasis was turned on during this acquisition. Figure 5.7 contains maps of the relative anisotropy [19] computed in three representative slices from uncorrected images (top row) and images that were coregistered using our algorithm (bottom row). The relative anisotropy is a diffusion anisotropy index that corresponds to the coefficient of variation of the three eigenvalues of the diffusion tensor in each voxel [19]. In the anisotropy maps computed from the uncorrected images several white-matter structures that are clearly visible in the anisotropy maps computed from the registered images cannot be identified. This is most evident in frontal regions where motion was most severe. In addition, the anisotropy maps computed from the uncorrected images show an artifactual rim of increased anisotropy around the periphery of the brain.

For the same dataset, Figure 5.8 shows representative slices of the $\chi^2$ maps produced by the tensor fitting procedure. The top row contains the $\chi^2$ maps of the original data prior to correction. The bottom row shows the $\chi^2$ maps after correction. For display purposes, in order to increase the dynamic range of the displayed image and reveal more inner structure, the square root of $\chi^2$ is shown. In this experiment, the mean $\chi^2$ of the image was reduced by about 80 percent after correction with our approach. One can note that the $\chi^2$ maps corrected datasets still contain slight bands of increased $\chi^2$ in some regions of the brain. One possible explanation for the origin of these bands is that we did not remove the volumes acquired during the voluntary motion. In these volumes there is significant intra-volume misregistration that our program does not correct for and which could account for these residual artifacts.

For the same dataset, Figure 5.9 shows the coefficients for the first 16 prin-

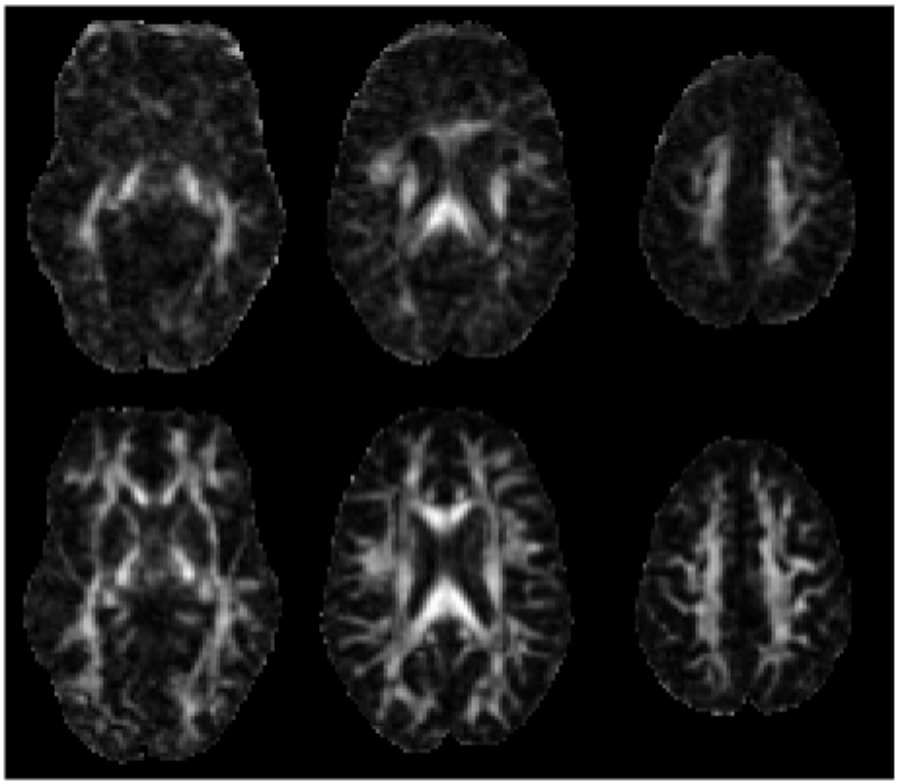

Figure 5.7: Axial views of the relative anisotropy index for three representative slices. Top row: anisotropy indexes computed from DWIs without alignment. Bottom row: anisotropy indexes computed from DWIs corrected for alignment. Apparent anisotropy around the top edges of the images seems to be significantly reduced in the corrected images.

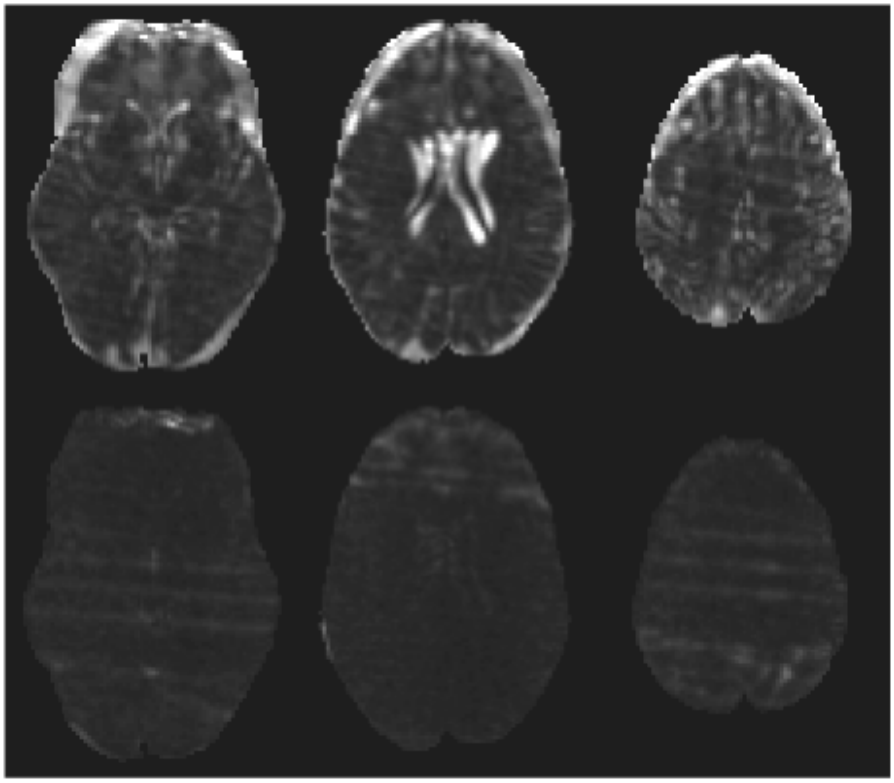

Figure 5.8: Axial views of the fit of the diffusion tensor for three representative slices. Top row: fit without alignment. Bottom row: fit from DWIs corrected for alignment. The chi squared error between the ADT model and the DWI data is significantly reduced after motion and distortion correction, indicating better image alignment.

cipal components in each voxel for a DWI dataset corrected for misalignment. The coefficients of the components are displayed in order of decreasing variance, starting from the top left. The coefficients of the first two components shown in Figure 5.9 appear to be related to T2-weighted contrast from the different tissues and to signal attenuation due to isotropic diffusion. The coefficients of components 4 to 8 clearly show effects related to anisotropic diffusion in white matter. Components 9 and higher show mostly noise. Figure 5.10 shows the same decomposition for DWIs that have not been co-registered. Not only do the first few components appear blurred, but most components higher than eight contain significant coefficient variability. For the PCA analysis of both registered and unregistered data, the volumes acquired when the voluntary motion occurred have been removed and striations such as those observed in Figure 5.8 are not visible. An interesting observation from the uncorrected dataset shown in Figure 5.10 is that the features related to diffusion anisotropy and motion induced artifacts do not separate into distinct components.

Figure 5.11 contains the plot of the relative variance of the coefficients of components 3 to 16 for the corrected and uncorrected datasets. Total variance (sum of the variances of each component) for a dataset was normalized to 1 for both datasets. The first two components account for 82 % and 91 % of the variance in the uncorrected and the corrected DWI dataset, respectively. The plot shows that fewer principal components are required to describe the intensity signal in the DWI dataset after motion and distortion correction. This confirms that fewer sources of signal are present in the data after correction. Given that no qualitative loss in signal anisotropy was detected after correction, we conclude that the effects of sources related to image misalignment have been reduced after

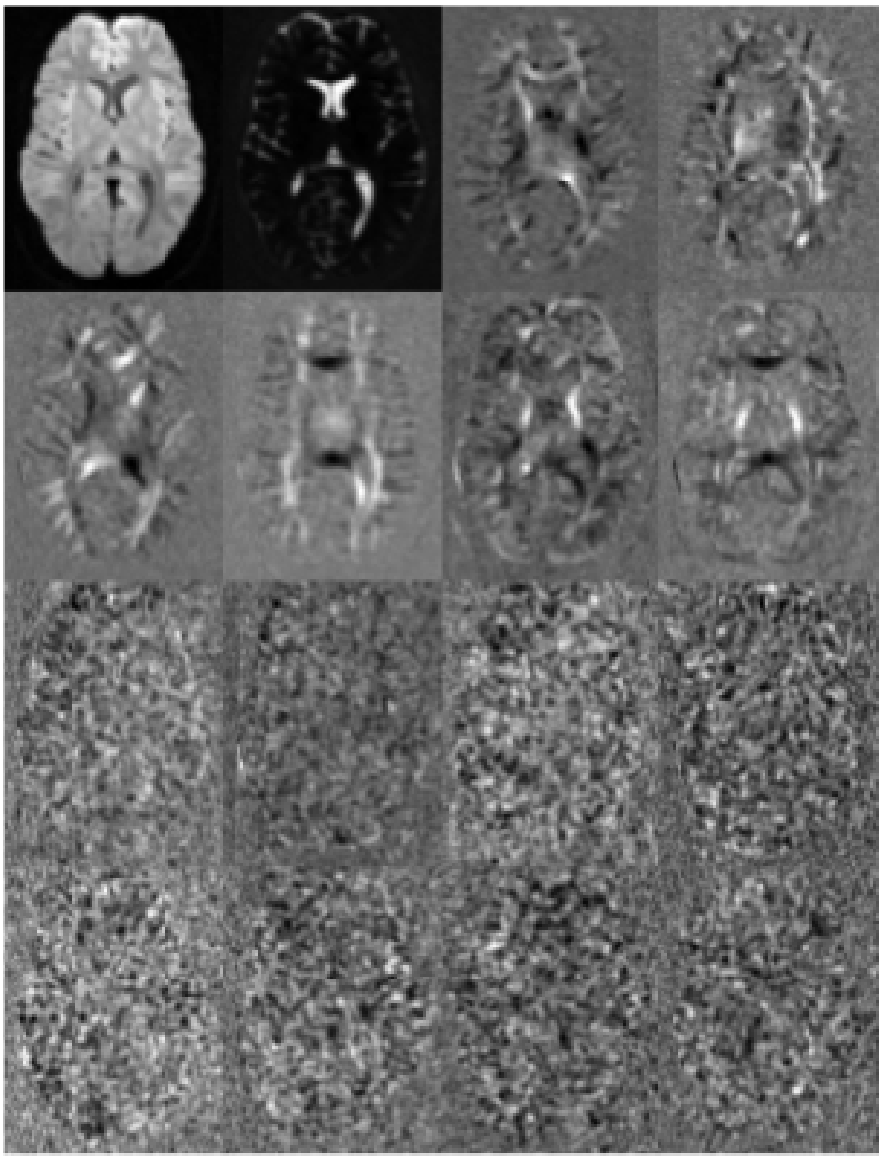

Figure 5.9: PCA decomposition of the registered DWI dataset (first 16 components). Images are displayed in order of decreasing variance, starting from the top left. Image variance around the edges of the brain, and in CSF-white matter interfaces, seems to be reduced in the third and fourth components.

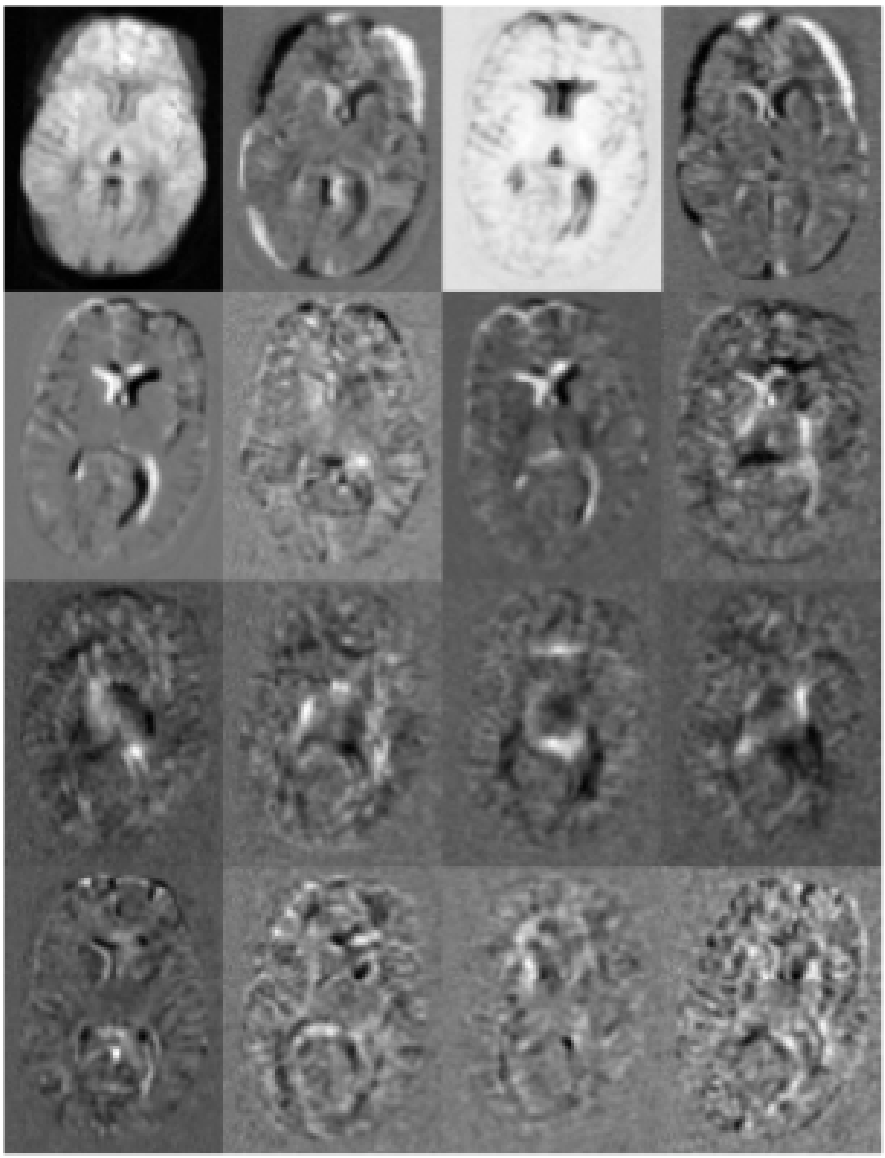

Figure 5.10: PCA decomposition of an unregistered DWI dataset (first 16 components). Images are displayed in order of decreasing variance, starting from the top left. Data variance around the edges of the images and CSF-white matter interfaces is apparent in nearly all components.

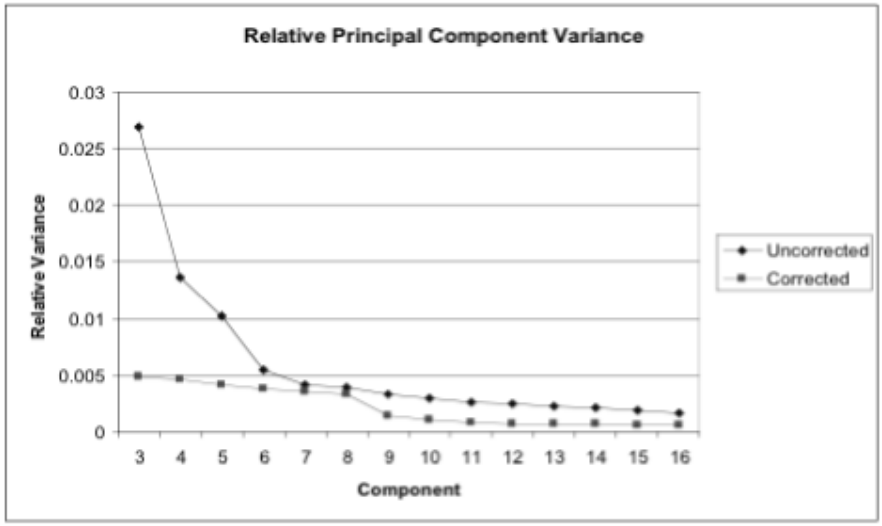


Figure 5.11: Relative variance of principal components, starting from the third component, for both registered and unregistered DWI datasets. The relative variances of the third and fourth components are reduced in the corrected dataset.

image registration.

## 5.4 Discussion

We have presented a novel method for correcting image distortion and for patient motion in DWI datasets. We use a mutual information-based registration algorithm to align each DWI volume in a dataset to a target volume chosen from the same DWI dataset. The registration is performed in 3D, with the warping function allowing for rigid body patient motion as well as eddy-current-induced distortion. All parameters are optimized simultaneously so that the final registration result represents an optimal correction of both patient motion and image distortion. After registration, the image intensity of each DWI volume is adjusted according to the spatial transformation applied to it. This prevents eddy-current-

induced distortions from introducing directional bias in the computed tensors. Similarly, each b-matrix is properly rotated using the same rotation applied to the corresponding DWI.

Amongst the novel aspects of our approach are using a model for eddy-current-induced fields based on Laplaces equation. A derivation of the image distortion caused by such fields as well as the brightness correction term necessary for adjusting the intensity values of the images is also provided. Lastly, if desired, the target volume for the DWI dataset can first be registered to a template. Subsequent registration of the DWI to the target T2-weighted volume will then cause the whole DWI dataset to be registered to a normalized template without additional computational cost, and with only one interpolation of the images. Thus our correction framework not only removes motion and distortion artifacts, but also positions the dataset in a standardized orientation using a single interpolation step. Note that, although we have designed our method with the purpose of spatially aligning DWIs of the human brain, it can be easily adapted to work with other types of images. The method is relatively fast. Our code written in IDL (Research Systems Inc.) aligns each 3D volume of size 128x128x72 in about 3 minutes on a Linux machine equipped with a 2 GHz Xeon processor. Implementing the method in a more efficient computer language, such as C, would certainly reduce the computation time for each image.

Results show a significant increase in data quality. Validation included visual inspection of the data as well as more quantitative measures such as the study of the $L_2$ error of the fitting of the data to the $\mathbf{D}$ model. In addition, PCA decompositions were used to study the data variance introduced by image misalignment. All results presented show that the quality of the DWIs datasets is significantly

improved after alignment. When large distortions are present, it is necessary to remap the intensity values of the aligned DWI according to the Jacobian matrix of the transformation to avoid directional bias artifacts. When relatively small distortions are present, such artifacts are difficult to detect visually, though the directional bias in the tensor field may still be present. The same can be expected for the b-matrix rotation step. If significant patient motion is present, rotating the b-matrix is essential to avoid erroneous computation of diffusion parameters.

Patient motion and eddy-current-induced image distortion are a common problem in clinical DWI acquisitions. Gradient pre-emphasis schemes that are now implemented in most MRI scanners are very effective in reducing the impact of eddy currents. In our scanner, eddy- current-induced distortions rarely exceed one or two pixels when pre-emphasis correction is applied. However, gradient pre-emphasis needs to be calibrated periodically. Datasets acquired immediately after calibration will have fewer artifacts than those acquired when a long period of time has elapsed since calibration. This temporal inconsistency in the quality of DWI data may be problematic in longitudinal studies and in general when the possibility of comparing scans acquired over time is desired. In our clinical studies the systematic use of our correction scheme has significantly increased the reproducibility of our clinical diffusion studies. In our experience, misalignment artifacts caused by patient motion are more problematic, especially in lengthy acquisitions where several DWIs are needed, or even in short scans with uncooperative patients or with unsedated pediatric subjects. Our results show that datasets containing significant motion can be successfully corrected.

Our DWI normalization method is still unable to correct some motion artifacts that can be present in DWI acquisitions. We perform a 3D registration between

brain volumes rather than a 2D registration between individual slices. Given that the 3D volumes are assembled from separate slice acquisitions there may be patient motion from one slice acquisition to another that will be uncorrected by our approach. We chose to perform a 3D registration because it is generally more robust than a 2D registration. Moreover, with a 2D registration, correcting for in-plane motion is feasible but correcting for out-of-plane motion is much more difficult. With our approach we can correct for some types of out-of-plane motion; for example, in axial images we can correct for translations in the z axis and rotations about the x and y axis. We also tested the possibility of performing the 3D registration first, followed by a 2D registration to correct for in-plane motion between slices, but we did not find that this strategy led to a significant improvement in the alignment of the images. Clearly the order of data acquisition is very important for a 3D approach to be effective: all slices composing a volume should be acquired in the shortest possible time. One additional problem is that the human brain is not strictly a rigid structure and cardiac induced motion may also deform the brain in ways that we are not currently able to account for with our model. Future work in the area could include characterizing the amount of image misalignment due to cardiac pulsation so that appropriate correction methods can be devised.

Finally, we anticipate that the algorithm presented here may not be able to register properly images acquired with diffusion weighting much higher than that used in this paper (b = 1100 s/mm$^2$). At very high b-values, the image signal in the brain parenchyma is significantly attenuated so that it may be confounded with signal from the air, nearly eliminating tissue/air boundaries that are some of the main features that guide the image registration process. Under these cir-

cumstances, the probability density functions (intrinsic contrast) of T2WI and DWIs is too dissimilar and even a mutual information-based registration algorithm would fail. One possible solution to this problem could be to perform image registration in a sequential or hierarchical manner. Rather than registering each heavily weighted DWI volume to a reference T2WI volume, one can obtain intermediate images that have more similar probability density functions and, therefore, are easier to register. For example, replicate volumes acquired with the same strength and orientation of the diffusion gradients can first be registered among themselves (with a rigid body transformation) and averaged to increase signal to noise. The next step would be to obtain trace-weighted volumes with improved signal to noise and anisotropy information removed. Compared to the original DWI volumes, trace-weighted volumes will have a probability density function more similar to that of T2 weighted volumes. Trace-weighted volumes can be obtained by computing the geometric average of a set of volumes acquired using b-matrices that must have certain properties (see [16] for more details on how to compute a trace-weighted DWI). Volumes obtained from the first averaging step can be registered among themselves and geometrically averaged. The resulting trace-weighted volume with the lowest b-value can in turn be registered to the T2WI volume, becoming the reference image for the trace weighted volume with the second lowest b-value. This process can be repeated to register all trace weighted volumes up to the highest b-value, achieving the goal of always registering volumes with similar contrast. The various transformations involved in this process can be combined and applied to each original DWI volume, avoiding artifacts originating from sequential interpolations. We are currently testing this approach and we hope that it will extend the ability of our method to register

heavily diffusion weighted images.

## 5.5 Appendix

### 5.5.1 Derivation of equation 5.8

Assuming that $b_{\text{eddy}}(\mathbf{y}_\alpha, t)$ is approximately constant during the readout and in steady state for the acquisition of different slices, the pixel values for volume $\alpha$ are

$$\begin{aligned} I_\alpha(m,l,q) &= \sum_{n=-\frac{N}{2}}^{n=\frac{N}{2}-1} \exp\left(-\imath\frac{2\pi mn}{N}\right) \sum_{p=-\frac{P}{2}}^{p=\frac{P}{2}-1} \exp\left(-\imath\frac{2\pi lp}{P}\right) \\ &\quad \times \int_{-\infty}^{\infty} dy_{\alpha 1} \int_{-\infty}^{\infty} dy_{\alpha 2} W(y_{\alpha 1}, y_{\alpha 2}, q s_{\alpha 3}) \exp\left(\imath(\phi_1 + \phi_2)\right) \end{aligned} \tag{5.23}$$

where $W$ the magnetization density of the object, $\phi_1 = \gamma G_1 y_{\alpha 1}$, $\phi_2 = \gamma(G_2 y_{\alpha 2} + b_{\text{eddy}}(\mathbf{y}_\alpha)\tau_{\text{echo}})$, $G_1$ is the product of the readout gradient and the sample time, and $G_2$ is the area of the phase encode blip. The value of the double integral in equation (5.23) is the measurement raw data in slice $q$ for readout point $n$ and phase-encode $p$, and the double sum is the discrete Fourier transform in the reconstruction. Exchanging the order of the sums and the integrals and regrouping, equation (5.23) becomes

$$\begin{aligned} I_\alpha(m,l,q) &= \int_{-\infty}^{\infty} dy_{\alpha 1} \int_{-\infty}^{\infty} dy_{\alpha 2} W(y_{\alpha 1}, y_{\alpha 2}, q s_{\alpha 3}) \\ &\quad \sum_{n=-\frac{N}{2}}^{n=\frac{N}{2}-1} \exp\left(-\imath n \Phi_1\right) \sum_{p=-\frac{P}{2}}^{p=\frac{P}{2}-1} \exp\left(-\imath p \Phi_2\right) \end{aligned} \tag{5.24}$$

where $\Phi_1 = \gamma G_1(y_{\alpha 1} - m s_1)$, $s_1 = \frac{2\pi}{N\gamma G_1}$ is the pixel size in the $y_1$ direction, $\Phi_2 = \gamma G_2(y_{\alpha 2} + \beta b_{\text{eddy}}(\mathbf{y}_\alpha) - l s_2)$, $s_2 = \frac{2\pi}{P\gamma G_2}$ is the pixel size in the $y_2$ direction

and

$$\beta = \frac{\tau_{\text{echo}}}{G_2} \tag{5.25}$$

The sums in equation (5.24) can be evaluated analytically [3], transforming (5.24) into

$$I_\alpha(m,l,q) = \int_{-\infty}^{\infty} dy_{\alpha 1} \int_{-\infty}^{\infty} dy_{\alpha 2} W(y_{\alpha 1}, y_{\alpha 2}, q s_{\alpha 3}) H_N\left(\Phi_1\right) H_P\left(\Phi_2\right) \tag{5.26}$$

where

$$H_N(\Phi) = \exp\left(\imath \frac{\Phi}{2}\right) \frac{\sin(N\frac{\Phi}{2})}{\sin(\frac{\Phi}{2})} \tag{5.27}$$

$H(\Phi)$, the point spread function of the acquisition, is large only close to the points where the denominator vanishes. The contribution to the integral from the lines centered at $\Phi \neq 0$ is wrap-around due to aliasing and vanishes if the field of view is large enough. To evaluate the integrals in equation (5.26) we perform the change of variables

$$x_{\alpha 1} = y_{\alpha 1} - m s_1 \tag{5.28}$$

$$x_{\alpha 2} = y_{\alpha 2} + \beta b_{\text{eddy}}(\mathbf{y}_\alpha) - l s_2 \tag{5.29}$$

which transforms equation (5.26) into

$$\begin{aligned} I_\alpha(m,l,q) \;\; = \;\; & \int_{-\infty}^{\infty} dy_{\alpha 1} \int_{-\infty}^{\infty} dy_{\alpha 2} W(x_{\alpha 1} + m s_1, y_{\alpha 2}, q s_{\alpha 3}) J(\mathbf{x}_\alpha) \\ & \times H_N\left(\gamma G_1 x_{\alpha 1}\right) H_P\left(\gamma G_2 x_{\alpha 2}\right) \end{aligned} \tag{5.30}$$

where $J(\mathbf{x}_\alpha) = \left|\frac{\partial(y_{\alpha 1}, y_{\alpha 2})}{\partial(x_{\alpha 1}, x_{\alpha 2})}\right|$ is the Jacobian determinant of the inverse of transformation (5.28,5.29), and $y_{\alpha 2}(\mathbf{x}_\alpha)$ is computed from equation (5.30). If we view $H(\Phi)$ as an approximation of a Dirac delta function, equation (5.30) tells us that the measured image $I_\alpha(m,l,q)$ is the true image shifted in the $x_2$-direction and weighted by the Jacobian determinant. If the distortion is small compared to the

distance over which $b_{\text{eddy}}$ changes appreciably, the inverse of equation (5.29) can be approximated as

$$y_{\alpha 2} = x_{\alpha 2} - \beta b_{\text{eddy}}(\mathbf{x}_\alpha) + ls_2 \tag{5.31}$$

with

$$b_{\text{eddy}}(\mathbf{x}_\alpha) \cong b_{\text{eddy}}(\mathbf{y}_\alpha). \tag{5.32}$$

The Jacobian then becomes

$$J = 1 + \frac{\partial(\beta b_{\text{eddy}})}{\partial x_2} \cong 1 + \frac{\partial(\beta b_{\text{eddy}})}{\partial y_2} \tag{5.33}$$

and equation (5.30) can be written as

$$I_\alpha(m, l, q) = \tilde{W}_\alpha \left( ms_1, ls_2 - \beta b_{\text{eddy}}(ms_1, ls_2, qs_3), qs_3 \right) \tag{5.34}$$

where the function $\tilde{W}_\alpha$ is the undistorted image of the volume smoothed by the point spread function.

# Chapter 6

# Correction of Motion Artifact in Cardiac Optical Mapping Using Image Registration

## 6.1 Introduction

Optical recording techniques have been widely employed in cardiac electrophysiology for studies of electrodynamics. Optical mapping is based on the proportional change of the induced fluorescence intensity resulting from the change in the transmembrane potentials in dye-stained tissue. The most significant constraint in cardiac optical recording is muscle contraction, which alters the fluorescence intensity and deforms the shape of the optical potentials. When the tissue moves during the recording, its relative location to the sensor and the light source changes, resulting in an artificial variation of fluorescence intensity intermingled with the desired signal. Most significantly, quantification of intensity variation is not meaningful if the fluorescence is recorded from different sites on the tissue in the same recording episode. Tissue contraction starts immediately after the upstroke of action potential. Therefore motion artifacts are more pronounced during action potential plateau when contraction is maximal and during repolarization phase when relaxation occurs. As a consequence, the correct mea-

surements of many interesting electrophysiological phenomena, such as action potential duration (APD) and repolarization, become impossible.

Common approaches to dealing with motion artifacts in fluorescence recording include mechanical constraining and chemical immobilization methods [32, 63, 41, 67] which are applied prior to imaging. In this work, we instead propose a retrospective motion correction approach that is based on a post-processing software technique known as image registration to spatially align the sequence of digital images taken from the optical recorder such that each location in the images acquired represents an intensity measurement of the same tissue location throughout the recording episode. We have adapted an existing registration technique used in medical imaging (see Maintz *et al.* [72] for an overview of medical image registration) so as to suit the epicardial fluorescence imaging data. Note that the work presented in this chapter was published, with modifications, in Rohde *et al.* [93].

## 6.2 Materials and Methods

### 6.2.1 Epifluorescence mapping

The experimental procedure was similar to that of a previous study [67]. In brief, New Zealand white rabbits weighing 4.4-5.5 kg were injected with 1000 units of heparin and 70 mg/kg sodium pentobarbital to induce deep general anesthesia. The heart was excised and the ascending aorta cannulated and secured for retrograde perfusion of the coronaries with a modified HEPES perfusate. The potential-sensitive dye di-4-ANEPPS (Molecular Probes, OR) at a concentration of 0.5 $\mu$M was added to the perfusate for approximately 15 minutes to stain the

heart. Fluorescence from the heart surface was elicited by a solid-state, frequency-doubled laser (Verdi V5, Coherent, Santa Clara, CA) at a wavelength of 532 nm. Laser light was delivered to the heart using multiple 1-mm optical fibers (SP-SF-960, FIS Inc., Oriskany, NY). The root-mean-square variation of laser intensity was 0.02 percent. The emitting fluorescence was imaged with a high-speed CCD camera (Model CA-D1-0128T, Dalsa Inc., Waterloo, ON, Canada) through a color glass filter with a cut-off wavelength of 600 nm (R60, Nikon).

### 6.2.2 Motion Correction via Image Registration

Given a series of $N$ digital frames from the experimental setup described above $\{I_1(\mathbf{x}), I_2(\mathbf{x}), \cdots, I_N(\mathbf{x})\}$ reference frame for all other images. We call this reference frame $T(\mathbf{x})$. We then proceed to align each frame $k$ in the original image sequence to the reference frame using the image registration algorithm proposed in [70], which finds the optimal alignment by solving numerically the following optimization problem:

$$\arg\max_{f_k} Q\left(I_k(f_k(\mathbf{x})), T(\mathbf{x})\right) \tag{6.1}$$

where $Q(\cdot,\cdot)$ is the Mutual Information image similarity measure and $f_k(\mathbf{x}) = A_k\mathbf{x} + \mathbf{t}_k$ is a 2D affine spatial transformation containing six independent parameters capable of performing rotation, scaling, shear and translation, with $\mathbf{t}_k = \{(tx)_k, (ty)_k\}$. Maximization of mutual information was first proposed in medical image registration problems in [70, 56] and has been shown to be robust in matching images whose intensity values are not linearly related [105, 121].

The mutual information similarity measure is given by

$$Q\left(I_k(f_k(\mathbf{x})), T(\mathbf{x})\right) = \sum_{i,t} p_{I_k,T}(i,t) \log\left(\frac{p_{I_k,T}(i,t)}{p_{I_k}(i)p_T(i)}\right) \tag{6.2}$$

where $p_{I_k}(i)$ and $p_T(t)$ are the marginal probability density functions (*pdf*) of $I_k(f_k(\mathbf{x}))$ and $T(\mathbf{x})$, respectively, and $p_{I_k,T}(i,t)$ is their joint *pdf*. The joint *pdf* $p_{I_k,T}(i,t)$ is computed from the normalized joint histogram of the images $I_k(f_k(\mathbf{x}))$ and $T(\mathbf{x})$ which is a matrix. The $(i,t)$ entry of this matrix stores the number of pixels that have intensity $i$ in image one and intensity $t$ in image two. When divided by the total number of pixels, it is an estimate of the joint probability function of the intensity values of the two images. Marginal distributions are obtained by summing along the lines and columns of this matrix. In this work, we have used 64 bins (i.e., we have divided the intensity range into 64 intervals) to create the joint histograms. Note that indexes $i$, $t$, for which $p_{I_k}(i)$, $p_T(t)$ or $p_{I_k,T}(i,t)$ is zero are not included in the computation of (6.2). Given a transformation $f_k(\mathbf{x})$ the image $I_k(f_k(\mathbf{x}))$ is computed using bilinear interpolation.

As in [70], we use Powells direction set method to compute the affine parameters that solve equation (6.1) [86]. Powells direction set method requires only the evaluation of function values for optimizing a cost function. It goes about finding a minimum by using a set of conjugate, or non-interfering, directions that are updated iteratively. For a quadratic cost function, it can be shown that Powells method finds the minimum of the function in $M(M+1)$ line minimizations, where $M$ is the number of parameters the cost function is dependent upon. The optimization is initialized with an identity transformation: all translation, rotation, and shear parameters set to zero, while the scaling parameters are set to 1. This ensures that $f_k(\mathbf{x}) = \mathbf{x}$. The tolerance value is $1.0 \times 10^{-4}$. Failure to increase the value of the similarity measure by more than $1.0 \times 10^{-4}$ in one optimization iteration signals completeness.

### 6.2.3 Data Analysis

Preliminary analysis of the data was done visually, focusing on regions of interest in which heart motion was evident. We analyzed the effect of our registration method on tissue activation quantitative measures such as action potential duration (ADP), activation isochrones, as well as abnormal or excessive depolarization and repolarization extracted from the movie sequences. The tissue activation timing of each pixel is detected by the peak of the first derivative of the time variation of the recorded fluorescence intensity. Activation isochrones were constructed from all the activation timing in the entire image. The frame of tissue at diastole ($\mathbf{F}_{rest}$) was defined as the frame right before the activation wavefront entered the field of view, whereas the frame with peak transmembrane potential ($\mathbf{F}_{peak}$) was selected when the entire field was depolarized. Because the fluorescence intensity is negatively proportional to the amplitude of transmembrane potential, $-(\mathbf{F}_{peak} - \mathbf{F}_{rest})$ represents the peak amplitude of the optical transmembrane potential ($\mathbf{F}_{amp}$). For every pixel in the image sequence, the "excessive depolarization" was calculated from the maximum deviation from $\mathbf{F}_{peak}$ in the depolarized direction. Similarly, the "excessive repolarization" was calculated from the maximum deviation from Frest in the repolarized direction. These two quantities were represented as a percentage of ($\mathbf{F}_{amp}$) for all the pixels. The action potential duration (APD50) was measured from the activation to the time when the amplitude fall below 50 percent of Famp in the repolarizing phase. Lastly, positive and negative deflection artifacts are also visible in the individual traces (image intensity of a fixed pixel location over time) of the movies. We also include sample traces obtained before and after correction with our method.

## 6.3 Results

Figure 6.1 displays representative registration results produced by the method described herein. Sequences of frames before (part a) and after (part b) registrations are shown. To help elucidate the motion present in the original sequence and the improvements in the corrected sequence we have placed a marker at the same pixel location for all frames. The reduced relative displacements between anatomical features and the marker after our motion correction approach indicate good overall alignment between the different movie frames. Figure 6.2 shows the measured excessive repolarization (negative deflection) and excessive depolarization (positive deflection) before (row a) and after motion correction (row b). These effects have been greatly reduced after correction. The pattern of the activation potential wavefront propagation, as displayed by the movies isochrones, remains intact. In addition, the APDs become significantly more evenly distributed after registration. Note that in this mode of motion, the tissue did not move out of the imaging field significantly. Figure 6.3 shows two pairs of the original (top) and the corrected (bottom) traces. Pair A shows an upward deflection after the activation due to the motion, whereas pair B shows a downward deflection. Both these deflections are corrected using the image registration algorithm. Note that figures 1-3 show representative analyses results. The same experiment was conducted on a set of 11 movies. The results generated on these were similar to those presented here.

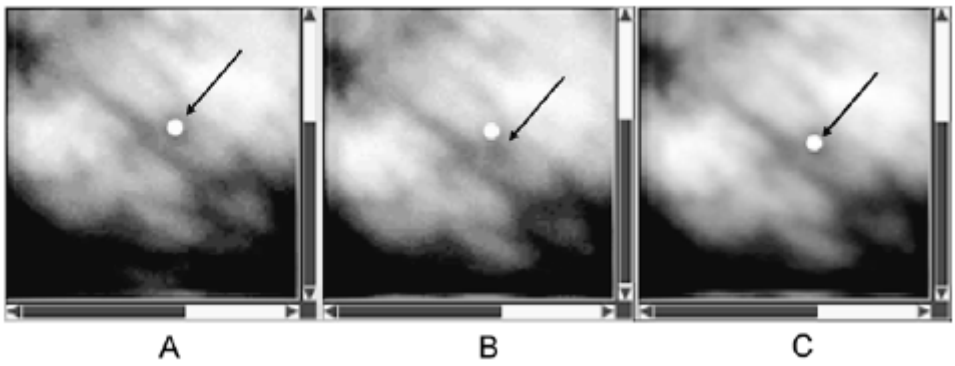


Figure 6.1: Images from a movie sequence: (A) reference frame; (B) maximum displacement before motion correction; (C) maximum displacement after motion correction. The relative displacement between the marker (white dot), which is placed at the same exact coordinate in all images, and image features (pointed by black arrow) is visibly reduced after motion correction.

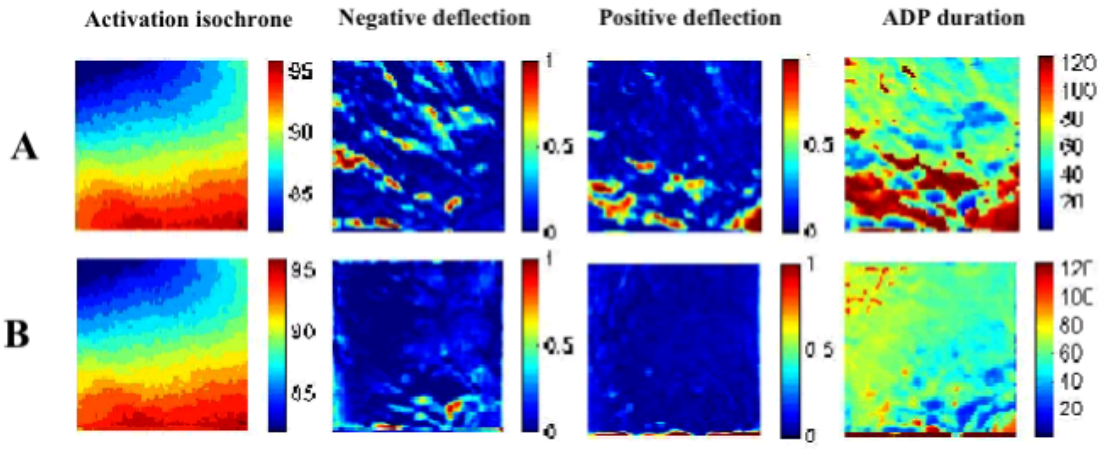


Figure 6.2: Activation isochrones, negative and positive deflection, and activation potential duration before (A) and after (B) correction. Activation isochrones remain relatively intact after motion correction while positive and negative deflection artifacts are reduced. Activation potential duration measurements are also more uniform after motion correction.

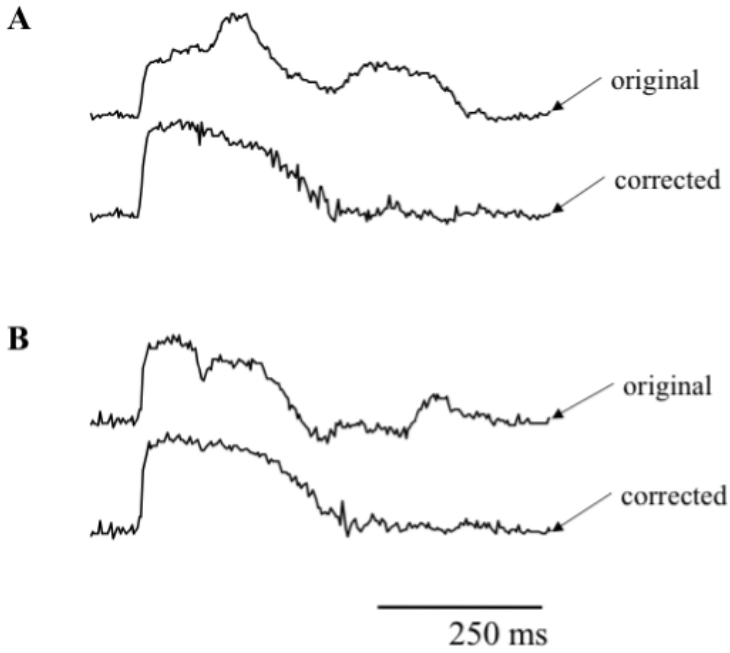


Figure 6.3: Two examples of trace extracted from original and corrected movies. Deflection artifacts are visibly reduced after motion correction.

## 6.4 Discussion and Conclusions

We have presented a simple method to correct for global motion present in epicardial fluorescence imaging experiments. Results showed that our approach significantly reduces motion artifacts of such image sequences. The algorithm is capable of reducing excessive depolarization and repolarization artifacts while preserving activation potential propagation. Activation potential duration is also more evenly distributed after correction with our approach. Our software was implemented in the IDL language (Research Systems, Inc.). The registration of each movie frame (128x128 pixels) takes about 5 seconds on a Pentium system running at 1.3 MHz. Implementing the software in a more efficient computer language such as C would certainly decrease computation time.

Note that the motion correction scheme presented here is one of many options. That is, instead of registering each movie frame to a single reference frame it is also possible to register each frame to the previous one. The advantage of the second approach is that the two consecutive frames should be more similar to each other. We have tried this option and, in our experience, this strategy is less stable than our current option. We have observed that while some corrections are good, some experiments generated results that were noticeably incorrect. We believe that the cause for this is the accumulation of successive registration errors. That is, the error for the registration of say frames 0 and 1 may be small, and so may be the error in the registration of frames 1 and 2. The error for the registration of frames 0 and 2, however, should in theory be larger than the error between 0 and 1 or the error between 1 and 2. Since the movie sequences we are using contain hundreds of frames, it is easy to see how this strategy can potentially generate highly inaccurate results. Yet another option is to use a single reference frame and initialize the optimization procedure using the result of the registration of the previous frame.

It is important to clarify the limitations of our approach. Firstly, all motions are assumed to be in plane. Thus out of plane motions cannot be corrected and in some instances may confound our correction approach. Furthermore, pixels that moved out of the imaging field due to motion cannot be recovered. Possible solutions to both problems include imaging the surface of the heart in stereo with multiple cameras or mirrors. This would allow collection of three-dimensional information that could be used to correct for motion artifacts out of the imaging plane. In addition, we only corrected for global motions (rotations, translations, shear and scaling) in this initial attempt to use image registration to correct for

motion artifacts. Local motions can be further corrected using nonrigid (nonlinear) registration methods. Preliminary results indicate that this is a promising direction, though technical implementation details can be complicated. In our experience we have found that the method presented above works best with images that have a small field of view focused on the surface of the heart. This could be related to the fact that while the affine spatial transformation model we use may be appropriate to describe local movement, it is not an appropriate model to simultaneously describe movement of several regions of the heart. The images presented in this paper have a field of view of about 20x20 millimeters, with resolution of about 200x200 microns.

In addition, it should also be noted that the method described above may fail to correct for motion in image sequences that have a large activation signal to noise ratio. In such cases, the algorithm may confound image features with activation signal, making motion correction difficult. The activation signal to noise ratio for our images falls typically in the 5 to 10 range. Initial experiments show that the algorithm works well for such images. At this point, however, we have not performed experiments to determine exactly at which activation signal to noise the algorithm starts failing.

# Chapter 7

# Estimating Intensity Variance Due to Noise in Registered Images: Applications to Diffusion Tensor MRI

## 7.1 Introduction

Post-acquisition image alignment (registration) is routinely performed in biomedical research and clinical practice [72, 84]. Applications using image registration techniques include motion and distortion correction in functional MRI (fMRI), diffusion tensor MRI (DT-MRI), and MR relaxometry experiments. In addition, image registration procedures are increasingly being used in computational based studies of neuroanatomy. This involves understanding the variability of tissue properties, including shape, across specific populations. An example is voxel-based morphometry, described in [11].

In general, many of the current post-processing methodologies can be summarized as follows. A set of medical images is acquired and reconstructed using standard methodologies. This step may include: filtering to avoid ringing artifacts, denoising, intensity corrections, etc. Next, using one of many available algorithms, images are registered to ensure, as much as possible, that a fixed image coordinate corresponds to the same structure, or anatomical coordinate, in

all images acquired. This step is necessary because the subject being imaged may move during data acquisition. In addition, images may contain geometric distortions with respect to each other. In echo planar (EPI) MRI these distortions can be caused by magnetic field susceptibility related artifacts. In EPI-based diffusion weighted imaging, significant geometric distortions may also occur due to eddy-currents induced by the rapidly switched diffusion weighting magnetic field gradients applied during imaging. Corrections to account for such misregistration artifacts are absolutely necessary to ensure the data analysis is reliable. In addition to correcting for motion and geometric distortions, the entire image sequence may also be aligned to a standard template image, using stereotaxic normalization techniques, for example, so that the data analysis results can be more conveniently interpreted. Data analysis consists of extracting or estimating some physically meaningful parameters from the sequence of medical images. In DT-MRI a 3x3 symmetric diffusion tensor is estimated, based on which several other quantities such as measures of diffusion anisotropy and depictions of fiber tracts can be generated. In fMRI, these may be statistical parametric maps [37], for example.

In many of these applications the analysis of the registered images involves fitting or estimating model parameters from the intensity values of the images. For such tasks it is crucial to know the correct signal variance of the registered images so that least-squares procedures, for example, can be properly implemented. Though significant research has been devoted to estimating signal variance in medical images–some examples in MRI include [49, 43, 100] among others–it is important to recognize that the signal variances in the registered and the original unregistered images differ. This is because the image interpolation or approxi-

mation step generally required in image registration can, as will be shown later, significantly change the noise properties of the image. We will show how a simple formula can be used to compute the appropriate signal variance in registered images. The analysis of diffusion weighted MRI data using the diffusion tensor model will be used as a case study. That is, given a set of diffusion weighted MR images (DWI) we use an existing software to register the DWIs to remove rigid body motion and eddy-current related distortions prior to tensor computation. We then show that noise variance in the registered images differs from the noise variance in the original images. However, even though DT-MRI is the only application discussed in detail in this chapter, we believe that the general approach described in this chapter should be considered whenever registered images are being analyzed using procedures that require knowledge of the variance in the image intensity values.

At the time of writing not much related work can be found in the biomedical imaging literature. Friston *et al.* [38] address the problem of removing movement-related artifacts, such as those caused by intensity fluctuations due to the change in position of the imaged object with respect to the reference frame of the scanner. In [106] and [42] the authors investigate the error in the intensity values produced by interpolation procedures applied on the registered images. Maas and Renshaw [69] discuss artifacts related to high frequency losses on registered (interpolated) data. Pluim et al. report that interpolation methods may cause undesirable artifacts when estimating the Mutual Information similarity measure [85]. Nickerson *et al.* [79] describe a method through which the local intensity variance in positron emission tomography (PET) can be estimated from the operations performed during image reconstruction. None of these works, however,

detail the importance of, and methods for obtaining correct estimates of the signal variance at each coordinate of each registered image.

In the field of diffusion weighted imaging and diffusion tensor MRI, several researchers have investigated methods for performing post-acquisition motion and distortion correction of data [47, 20, 55, 8, 73, 91]. Though the registration methods differ, most of these works use linear interpolation to produce the series of DWIs. This series is then used to estimate one diffusion tensor for each voxel via least- squares fitting procedures similar to the $\chi^2$ minimization procedure described in [17]. We show in this chapter that least-squares fitting procedures that extract diffusion tensor estimates from registered data can be affected by the changes in image noise properties due to interpolation. We also provide a simple method for obtaining correct variance estimates for the registered images. The work presented in this chapter was previously published, with some modifications, in Rohde *et al.* [92].

## 7.2 Theory

In practice, the process of registering two images is usually approached within an optimization framework in which the goal is to find a spatial transformation $f(\mathbf{x})$, where $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , or $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ for volumetric images, that maximizes some similarity measure $I$ between the digitized target $T(\mathbf{x})$ and source $S(\mathbf{x})$ images:

$$\max_{f} I\left(S(f(\mathbf{x}), T(\mathbf{x})\right) \tag{7.1}$$

The function $f(\mathbf{x})$ may be a rigid body, affine, or higher order transformation, depending on the application. The function $I$ usually measures the similarity between the images being registered by computing some form of statistical de-

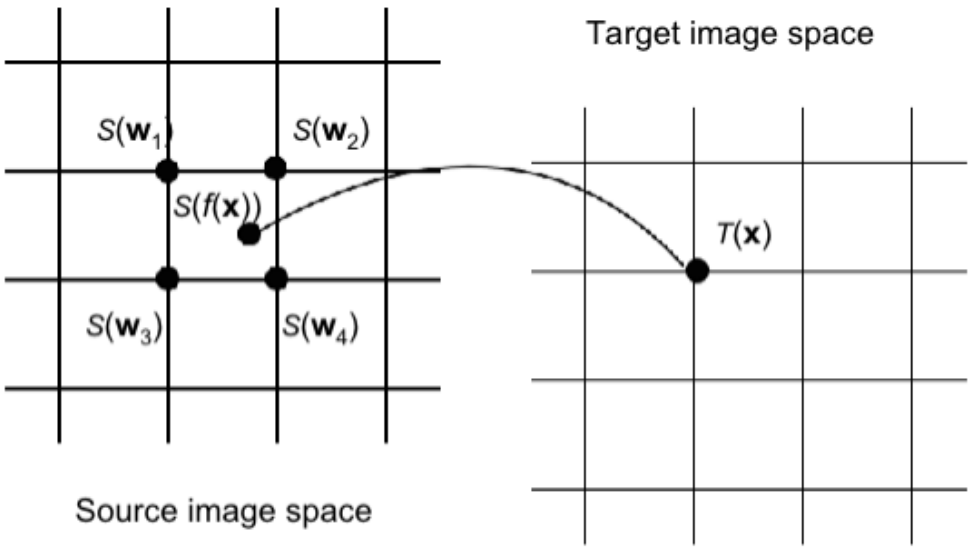


Figure 7.1: Illustration of an interpolation or approximation procedure for image registration. First, a coordinate $\mathbf{x}$ in the target image space is transferred to a coordinate in the source image space via $f(\mathbf{x})$. The value of the source image at $f(\mathbf{x})$ is computed using neighboring values of the source image at that coordinate, $s(\mathbf{w}_1)$, $s(\mathbf{w}_2)$, etc.

pendency between the intensity values of the images. In the processing pipeline described above, the problem defined by equation (7.1) is usually solved for $K$ images in the image sequence $\{S_1(\mathbf{x}), \cdots, S_K(\mathbf{x})\}$, so it is clear that the sequence of images $\{S_1(f_1(\mathbf{x})), S_k(f_k(\mathbf{x})), , S_K(f_K(\mathbf{x}))\}$ is properly aligned. Note that in cases where $f_k(\mathbf{x})$ is used to correct for geometric distortions caused by imperfect magnetic field gradients in MRI, for example, the intensity value of the corrected images may also have to be multiplied by a correction factor [103, 91]:

$$\tilde{S}_k(f_k(\mathbf{x})) = S_k(f_k(\mathbf{x})) \det |\mathrm{Jac}(f_k(\mathbf{x}))| \tag{7.2}$$

where $\det |\mathrm{Jac}(f_k(\mathbf{x}))|$ stands for the determinant of the Jacobian matrix of the transformation $f_k(\mathbf{x})$.

Independently of how the solution to equation (7.1) is actually computed for each image in the sequence, many imaging applications require knowing the value of the registered images $\{S_1(f_1(\mathbf{x}_i)), \cdots, S_K(f_K(\mathbf{x}_i))\}$ for some arbitrary coordinate $\mathbf{x}_i$. Since in general the point $f_k(\mathbf{x})$ will not coincide with a sampling coordinate of image $S_k$, an interpolation or approximation strategy must be used to produce the image value $S_k(f_k(\mathbf{x}))$. Many approximation and interpolation methods can be chosen to perform such tasks [75]. Most estimate the value of $S_k(f_k(\mathbf{x}))$ based on a linear combination of the intensity values of image $S_k$ around the point $f_k(\mathbf{x})$. Figure 7.1 illustrates this process. Note that $\mathbf{w}$ refers to grid coordinates of the image $S_k$. Mathematically, this interpolation or approximation procedure can be expressed as:

$$S_k(f_k(\mathbf{x})) = \sum_{\mathbf{w}_i \in \Omega} \alpha_i S_k(\mathbf{w_i}) \tag{7.3}$$

where $\Theta$ defines a set of sampling coordinates that surround $f(\mathbf{x})$ (see Figure 7.2). The coefficients $\alpha_i$ of the linear combination (7.3), as well as the size of $\Theta$ are determined solely by the choice of interpolation or approximation kernel. For the linear interpolation method, one of the most popular image interpolation methods, the value of the image $S$ at coordinate $f(\mathbf{x})$ is given by:

$$S(f(\mathbf{x})) = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}(1-V_i)(1-P_j)(1-Q_k)S(x_i, y_j, z_k) \tag{7.4}$$

where $V_i = |f(\mathbf{x}_x) - x_i|$, $P_j = |f(\mathbf{x}_y) - y_j|$, $Q_k = |f(\mathbf{x}_z) - z_k|$, and $\{x_i, y_j, z_k\}$ are image grid coordinates for which $|f(\mathbf{x})_x - x_i| < 1$, $|f(\mathbf{x})_y - y_i| < 1$, $|f(\mathbf{x})_z - z_i| < 1$. Thus the coefficients of the linear combination (7.3) are given by:

$$\alpha_{i,j,k} = (1-V_i)(1-P_j)(1-Q_k) \tag{7.5}$$

The set $\Theta$, in this case, are the coordinates $\mathbf{w}_i$ for which $|f(\mathbf{x}) - \mathbf{w}_i| \leqslant 1$ holds.

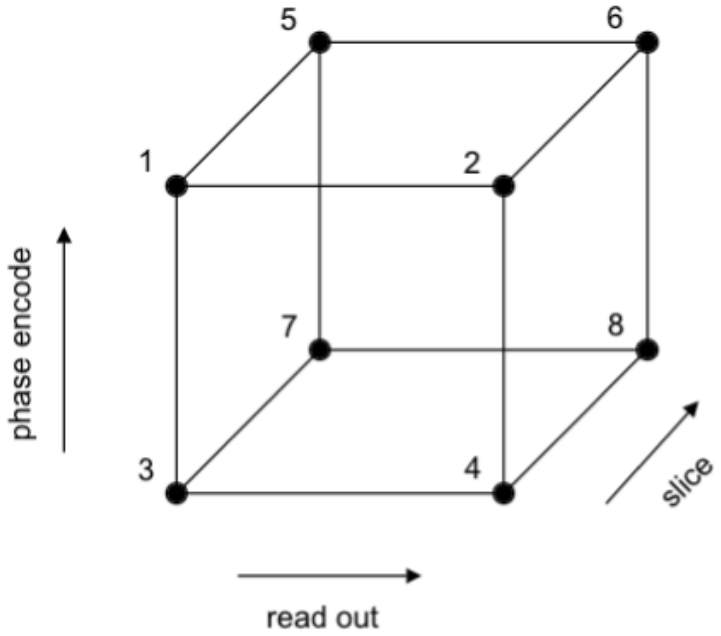


Figure 7.2: Ordering of voxels used to compute the correlation matrix 7.9. Corr(1,3), for example, corresponds to the correlation coefficient between the image value at location with index 1, and the image value at location with index 3.

Note that formulas (7.4), and (7.5) represent the three dimensional case, while figure 7.1 depicts a 2 dimensional situation. Naturally, when different interpolation or approximation methods are used, different formulas are needed for estimating the variance of any given interpolated image value. Refer to appendix A for the general formula for the variance of an interpolated image value given a general (separable) basis function.

Because of random variability introduced at several steps during image acquisition, the measurement $S_k(\mathbf{w}_i)$ should be considered a random variable with a variance $\mathrm{Var}\,(S_k(\mathbf{w}_i))$. For MR images it is customary to assume that noise variance, denoted by $\lambda^2$, is uniform throughout the imaging volume. Note that,

though it can be assumed that $S_k(\mathbf{w}_i)$ and $S_k(\mathbf{w}_j)$, where $i \neq j$, have equal variances, in general they are not independent measurements because several image reconstruction steps effectively correlate measurements from different image coordinates. Correlation in the data due to the reconstruction procedure can arise from filtering during analog to digital conversion, filtering to remove ringing artifacts, filtering to remove noise, correcting for ghosting artifacts (particularly salient in EPI reconstructions), and others. Correlation between values in different image coordinates occurs not only in MRI, but X-ray based computed tomography and positron emission tomography (PET) also. This is because most reconstruction algorithms use filtering operations that correlate intensity values of different image coordinates. A simple method for estimating this correlation in MRI will be described in the next section.

In short, because of the noise variability introduced during image acquisition and processing, the measurements $S_k(\mathbf{w}_i)$ and $S_k(\mathbf{w}_j)$ are random variables with variance $\mathrm{Var}\left(S_k(\mathbf{w}_i)\right)$ and $\mathrm{Var}\left(S_k(\mathbf{w}_j)\right)$, respectively, and covariance $\mathrm{Cov}\left(S_k(\mathbf{w}_i), S_k(\mathbf{w}_j)\right)$. Thus, $S_k(f_k(\mathbf{x}))$, as defined by equation (7.3), is also a random variable with variance [53]:

$$\begin{aligned} \mathrm{Var}\left(S_k(f_k(\mathbf{s}))\right) &= \left(\sum_{\mathbf{w}_i \in \Theta} \alpha_i^2 \mathrm{Var}\left(S_k(\mathbf{w}_i)\right)\right) + \\ & \quad 2\left(\sum_{\{\mathbf{w}_i, \mathbf{w}_j \in \Theta, i<j\}} \alpha_i \alpha_j \mathrm{Cov}\left(S_k(\mathbf{w}_i), S_k(\mathbf{w}_j)\right)\right) \end{aligned} \tag{7.6}$$

If it can be assumed that $\mathrm{Var}\left(S_k(\mathbf{w}_i)\right)$ is approximately constant for all values of

the image (7.6) simplifies to,

$$\begin{aligned} \mathrm{Var}\left(S_k(f_k(\mathbf{s}))\right) &= \lambda^2 \left( \sum_{\mathbf{w}_i \in \Theta} \alpha_i^2 \right) + \\ & \quad 2 \left( \sum_{\{\mathbf{w}_i, \mathbf{w}_j \in \Theta, i<j\}} \alpha_i \alpha_j \mathrm{Cov}\left(S_k(\mathbf{w}_i), S_k(\mathbf{w}_j)\right) \right) \end{aligned} \tag{7.7}$$

In cases when the intensity correction function defined in (7.2) needs to be applied to the registered image $S_k(f_k(\mathbf{x}))$ to obtain intensity corrected value $\bar{S}(f_k(\mathbf{x}))$, it is easy to show that the correct formula for the variance becomes:

$$\begin{aligned} \mathrm{Var}\left(S_k(f_k(\mathbf{s}))\right) &= \left(\det\left|\mathrm{Jac}(f_k(\mathbf{x}))\right|\right)^2 \lambda^2 \left( \sum_{\mathbf{w}_i \in \Theta} \alpha_i^2 \right) + \\ & \quad 2 \left(\det\left|\mathrm{Jac}(f_k(\mathbf{x}))\right|\right)^2 \left( \sum_{\{\mathbf{w}_i, \mathbf{w}_j \in \Theta, i<j\}} \alpha_i \alpha_j \mathrm{Cov}\left(S_k(\mathbf{w}_i), S_k(\mathbf{w}_j)\right) \right) \end{aligned} \tag{7.8}$$

Note that if nearest neighbor interpolation is used, the variance of each value in the interpolated image would be equal to the variance of the nearest neighbor voxel, multiplied by the Jacobian correction term when appropriate. The formula (7.8), using the linear interpolation method, was implemented in a typical DTI processing pipeline and it is shown next that, because the fitting procedure includes estimates of the noise variance in each image, formula (7.8) should be used to re-estimate the variance of the signal in each voxel in each image that has been registered.

## 7.3 Methods

### 7.3.1 MRI data acquisition

The data sets in the demonstrations used throughout this paper were acquired with a standard single-shot multi-slice spin-echo EPI sequence (i.e.: fat sup-

pression pulse, 90 degree pulse, first diffusion gradient, 180 degree pulse, second diffusion gradient, EPI readout). Scans were performed on a 1.5 T GE Signa system equipped with a whole-body gradient coil able to produce gradient pulses up to 50 mT/m (GE Medical Systems, Milwaukee, WI). The imaged volume was composed of 80 contiguous slices with 2 mm slice thickness and 2 mm in-plane resolution. The echo-time was 82.7 ms, the read-out time 50 ms, and the repetition time was greater than 10 s with cardiac gating (4 acquisitions per heart beat starting with a 150 ms delay after the rise of the sphygmic wave as measured with a peripheral pulse oxymeter). The gradient strength was 49 mT/m, yielding a b value (i.e., trace of the b-matrix) of 1,120 s/mm$^2$. A total of 56 3D images were acquired by repeating 8 times a diffusion sampling scheme described previously [83] which includes one volume with no diffusion weighting followed by the same volume six times, acquired with diffusion gradients applied in different directions. The total imaging time was approximately 20 minutes. Replicate volumes were acquired for signal to noise considerations in order to improve the quality of the estimated diffusion tensor parameters. The signal to noise ratio, as measured by the mean signal in the region of the thalamus divided by the estimated standard deviation of the signal (see section below), was about 13 for the T2-weighted images and about 7 for the diffusion weighted images.

### 7.3.2 MRI noise estimation

The sources that introduce uncertainty in each voxel intensity are many and are generally put into one of two categories: thermal noise, and physiological noise. Other sources may also exist in the electronics of the acquisition system, such as digitization etc., but these can be minimized in an ideal experiment. Thermal

noise is usually considered as white noise because it is expected that its power should be equal for all frequencies within the readout bandwidth. Because the images are reconstructed using the Fourier transform, the variance that characterizes the uncertainty due to thermal noise is constant throughout the imaging volume [44]. Naturally, the same cannot be said about physiological noise.

In our experiments we are only able to estimate the variance that characterizes the uncertainty of the MR measurement due to thermal noise. We do so by computing the variance of magnitude reconstructed intensity values in an artifact-free background region and propagating it to regions with strong signal from the brain through the method described in [49, 43]. The correction factor described in [49, 43] uses the assumption that Gaussian distributed noise is added to the real and imaginary channels of the receiver system. If possible, we would also like to estimate the variance component due to physiological noise such as flow, MR spin history errors, etc. To do so, however, would require many repeated acquisitions. It would also be difficult to isolate the variance due to patient motion in such repeated measurements (which is something the registration step is actually trying to diminish). Because of these difficulties, we are not able to estimate the variance introduced by physiological effects. Thus the variance estimate we are able to compute for each voxel intensity value is a biased lower bound estimate of the variance when all sources of uncertainty are included. We would like to note, however, that if the total variance, and covariance, in the signal (from all sources) does somehow become available in the future, the same formulas described in the theory section can be used to propagate the known variance beyond the interpolation step.

The correlation matrix used in our experiments was estimated empirically.

Though theoretically possible, it could be very cumbersome to account for all of the filtering steps applied to the data before it becomes a magnitude image. In addition, some steps taken during analog to digital conversion of the free induction decay signals may be proprietary and thus inaccessible. Instead, we acquired and reconstructed several 3D images of pure noise. Using this pure noise image data we computed the correlation coefficient between the original volumes and the same volumes shifted by one pixel in the x, y, and z directions. Note that because we are using linear interpolation, it is only necessary to include 1 voxel shift in the computation (7.8). When bases functions of wider support are used in the interpolation or approximation procedure, the correlations of larger shifts may be required. Using this method, we computed the following 8x8 correlation matrix:

$$\mathrm{Corr}(i,j) = \begin{pmatrix} 1 & 0.35 & 0.40 & 0.25 & 0 & 0 & 0 & 0 \\ 0.35 & 1 & 0.25 & 0.40 & 0 & 0 & 0 & 0 \\ 0.40 & 0.25 & 1 & 0.35 & 0 & 0 & 0 & 0 \\ 0.25 & 0.40 & 0.35 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0.35 & 0.40 & 0.25 \\ 0 & 0 & 0 & 0 & 0.35 & 1 & 0.25 & 0.40 \\ 0 & 0 & 0 & 0 & 0.40 & 0.25 & 1 & 0.35 \\ 0 & 0 & 0 & 0 & 0.25 & 0.40 & 0.35 & 1 \end{pmatrix} \tag{7.9}$$

Figure 7.2–which defines the ordering of the voxel coordinates–helps explain the correlation matrix expressed in (7.9). Because we are assuming that most of the correlation is caused by linear filtering operations applied on the image data, the noise correlation matrix (7.9) should be approximately constant throughout the domain of the original magnitude reconstructed images. Note that since our

acquisition is based on a 2D EPI pulse sequence, measurements between one slice and the next show no significant correlation. Also note that the correlations in the $x$, and $y$ directions are not equal, since additional operations are performed in the phase encode ($y$ in this case) direction to minimize ghosting artifacts. Lastly, since we are also assuming that the noise variance in the original magnitude reconstructed image is constant, the covariance matrix used in (7.9) is given by:

$$\mathrm{Cov}(i,j) = \mathrm{Corr}(i,j)\lambda^2. \tag{7.10}$$

### 7.3.3 Diffusion tensor estimation

The diffusion tensor model was estimated in each voxel $\mathbf{x}$ from the diffusion weighted data by minimizing the following equation:

$$\chi^2\left(\mathbf{D}(\mathbf{x}), A(\mathbf{x})\right) = \frac{1}{K-7}\sum_{k=1}^{K}\frac{\left(A(\mathbf{x})e^{-\mathbf{D}(\mathbf{x}):\mathbf{b}_k} - \tilde{S}_k(f_k(\mathbf{x}))\right)^2}{\mathrm{Var}\left(\tilde{S}_k(f_k(\mathbf{x}))\right)} \tag{7.11}$$

where $\mathbf{D}(\mathbf{x})$ is a 3x3 symmetric matrix, $A(\mathbf{x})$ is the amplitude term, and $\mathbf{b}_k$ is the b-matrix for image $k$, and $\mathbf{D} : \mathbf{b}$ stands for the matrix dot product ([17]. The minimization was performed using the Levenberg-Marquardt least-squares method.

### 7.3.4 Simulated data experiments

As an initial test of our variance estimation software we performed simulation experiments using artificially constructed data. In this experiment, one thousand 2D images of Gaussian distributed random noise with mean zero and variance one were rotated about their centers by 5 degrees using bilinear interpolation. In

this simulation, the correlation matrix used was approximately:

$$\mathrm{Corr}(i,j) = \begin{pmatrix} 1 & 0.25 & 0.25 & 0.23 \\ 0.25 & 1 & 0.23 & 0.25 \\ 0.25 & 0.23 & 1 & 0.25 \\ 0.23 & 0.25 & 0.25 & 1 \end{pmatrix}$$

For a fixed pixel coordinate $\mathbf{x}$ the variance across all of the rotated images was computed and displayed. The purpose of this experiment is to show that the variance in the images acquires a particular striped structure. The origin of the striped structure shown stems from the fact that each intensity value in the rotated image was computed by interpolating the original image on a particular non-grid-point coordinate. This estimate comes from a linear combination of the intensity values from around the transformed sampling coordinate (see fig. 1). The coefficients of the linear combination are computed from the distance of the transformed coordinates to its nearest neighbors. For a specific degree of rotation, this distance will repeat itself every so often throughout the rotated image domain. Since the variance of the rotated image is determined by the coefficients of the linear combination, the variance value of the rotated image will also repeat itself every so often throughout the rotated image domain. We show that by using equation (7.8), the variance in the interpolated images can be predicted exactly.

### 7.3.5 Experimental data

The diffusion weighted data used in the examples in this paper were registered to account for patient motion and eddy-current induced geometric distortions using the methodology described in [91]. When using this method, first a non diffusion

weighted image is chosen from the same DWI dataset to be the reference image to which all remaining images are aligned. The registration of each image is done in series, and independently from the registration of the other images in the same set. This approach uses a mutual information-based registration technique and a spatial transformation model containing parameters that correct for eddy-current-induced image distortion and rigid body motion in three dimensions. Each registration consists of estimating 14 parameters in total: 6 for rigid body motion, and 8 for the model of eddy-current induced distortions which consists of a spherical harmonics series expansion in Cartesian coordinates, up to quadratic terms. All 14 parameters for each image in the set are estimated simultaneously. Optimization is performed using a gradient-ascent-type technique within a multi-resolution framework. Initial estimates of the registration parameters are obtained using low-resolution approximations of the images. These estimates are then used to initialize the optimization using higher-resolution representations of the data. The images can also be registered to an arbitrary template with a single interpolation step without additional significant computational cost, though this feature was turned off in all of the experiments shown here. The registered images are created using trilinear interpolation. Following registration, the signal amplitude of each DWI volume is corrected to account for size variations of the object produced by the distortion correction, and the b-matrices are properly recalculated to account for any rotation applied during registration.

The diffusion tensor at each voxel was computed using the registered images by solving equation (7.11) as described above. For comparison purposes, we also estimate the diffusion tensor from the registered images using equation (7.11), but using a constant term for the noise variance $\mathrm{Var}(\tilde{S}_k(f_k(\mathbf{x}))) = \lambda^2$. The $\chi^2$

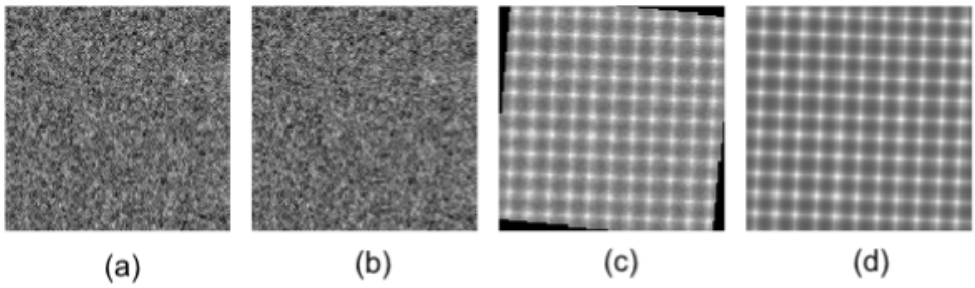


Figure 7.3: Simulation showing how the interpolation necessary to relate measurements in two images can significantly affect the noise properties of the interpolated image. Part (a): an image of simulated noise. Part (b) is the image in part (a) rotated by 5 degrees. Part (c) is the variance of image (b) computed by repeating the rotation experiment 1000 times. Part (d) shows the variance of image (b) predicted by formula (7.8).

measure at each voxel is compared for both methods. In addition to $\chi^2$ we also compare the estimated tensor parameters to investigate whether or not they are significantly affected when the incorrect noise variance is used.

## 7.4 Results

The results of the simulation experiments are shown in Figure 7.3. Part (a) shows a sample noisy image computed as described above. Part (b) shows the same image rotated by 5 degrees about its center. Values outside the original image were assumed to be zero. Part (c) shows an image of the variance of the one thousand rotated images computed at each pixel. Clearly the variance became non-uniform and acquired a striped pattern throughout the domain of the image. This variance image was computed analytically using formula (7.7), and the result is shown in part (d).

A similar effect can be seen in real data experiments using diffusion weighted images. Though these striped artifacts are practically invisible in the interpolated DWI volumes, they become evident in the $\chi^2$ maps computed using equation (7.11). Some results are shown in Figure 7.4. In this experiment a set of DWI volumes was rotated about its horizontal axis by about 7.5 degrees, thus causing interpolation to be performed between values of different slices, as well as between values of different lines in the logical $y$ direction. For this experiment, the same rotation transformation was applied to each DWI volume, that is: $f_1(\mathbf{x}) = f_k(\mathbf{x}) \cdots = f_K(\mathbf{x})$. Part (a) of Figure 7.4 shows the $\chi^2$ map computed using a single value, $\lambda^2$, for the variance of each voxel in each image. Horizontal stripes are visible along the vertical axis of the image, reflecting the different amounts of interpolation performed at each voxel location. Part (b) shows the variance predicted using eq. (7.8). Part (c) shows the $\chi^2$ map computed using the variance given by eq. (7.8). The striping patterns become negligible when compared to those shown in part (a) of the same figure. Note that the dynamic ranges of both $\chi^2$ maps in this example were auto-scaled to obtain maximum contrast.

Figure 7.5 displays an additional comparison of $\chi^2$ maps computed on unregistered and registered data, with and without the estimation of intensity variance method we propose above. Unlike the example above, the dynamic range of all $\chi^2$ maps was set to [0,5]. Note that the original image data used in this experiment was significantly misregistered due to relatively large subject motion. Part (a) shows the $\chi^2$ computed from the original, unregistered images using a single variance value estimated from the background of the images. Figure 7.5(b) shows the $\chi^2$ map computed from the registered DW images with the same variance value

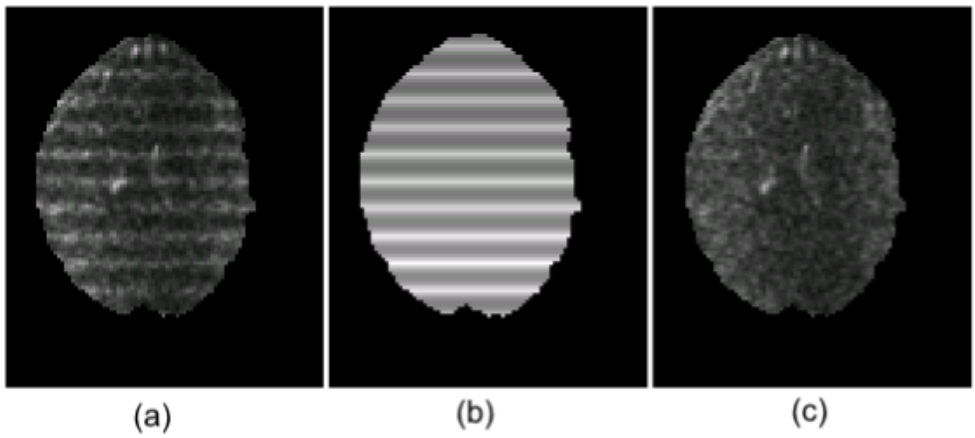


Figure 7.4: Demonstration of bias in $\chi^2$ between the DT model and registered DWI data. Part (a) show the $\chi^2$ map computed using a single value for the variance in the data. Part (b) shows the non-uniform variance estimated using formula (7.8). Part (c) shows the same $\chi^2$ map, however, this time computed using the variance values displayed in part (b).

used in part (a). Note that the $\chi^2$ values of the registered images are generally lower than the chi squared values of unregistered images. Finally, figure 7.5(c) shows the $\chi^2$ maps computed from the registered DW images using the variance values produced by eq. (7.8). Note also that the chi squared values for part (c) are generally higher than those of part (b).

We also compared some of the most well known parameters derived from the diffusion tensor computed from the fitting of eq. (7.11). For reference, the amplitude, trace, and fractional anisotropy index [14, 19] are shown in figure 7.6, parts (a) through (c), respectively. Figure 7.7(a) shows the relative error between the trace parameter computed with and without the variance correction scheme proposed above. The relative error was computed using the following formula: $|v_{corrected} - v_{uncorrected}|/v_{corrected}$, where $v$ stands for the voxels specific value for

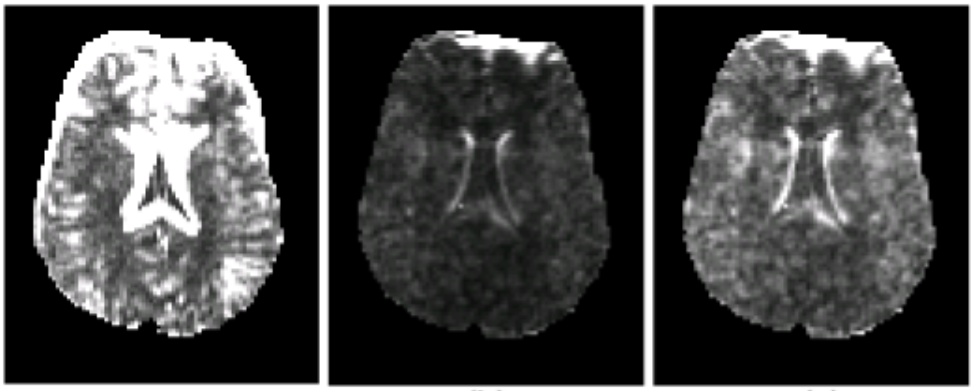


Figure 7.5: Part (a): $\chi^2$ images computed from the raw (unregistered) data. $\chi^2$ images computed from registered images with (part c) and without (part b) the noise variance formula given in eq. (7.8). The $\chi^2$ values computed using the correct noise variance values are generally higher than the values computed using a single noise variance estimated on the original (unregistered) images. The actual noise variance in registered images is generally lower than the original noise variance because of the linear combinations performed during image interpolation.

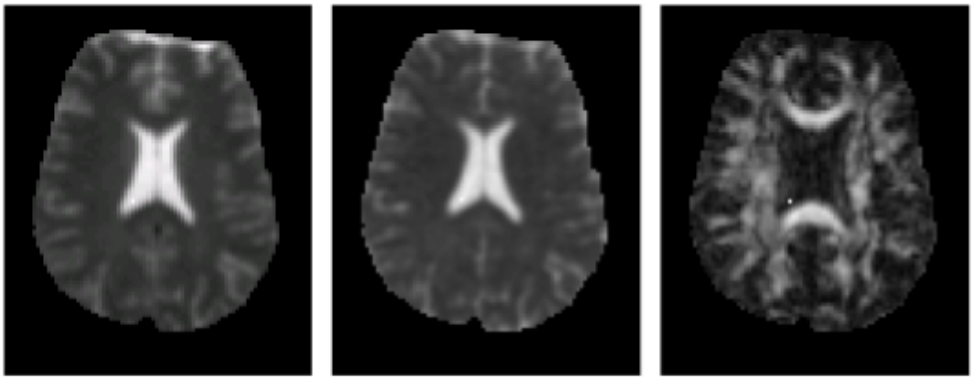


Figure 7.6: Tensor derived quantities computed after registration with correct variance estimates. Part (a): amplitude image. Part (b): trace of the diffusion tensor. Part (c): fractional anisotropy image.

the trace of the diffusion tensor. The absolute value of the difference between the fractional anisotropy values computed with and without the variance correction described above is shown in part (b) of figure 7.7.

## 7.5 Discussion

The rotation experiments performed with the simulated noisy images demonstrate qualitatively and quantitatively the effect that image interpolation can have on the noise variance in registered or interpolated imagesthe variance becomes non-uniform. The experiment also shows that formula (7.8) can be used to estimate the variance in the interpolated images.

Experiments using real DWIs showed that the change of image noise properties caused by the registration (interpolation) procedure can significantly affect parameter estimation procedure in DT-MRI. First, the alignment of the entire DWI dataset to a standard template can cause $\chi^2$ maps to acquire a striped

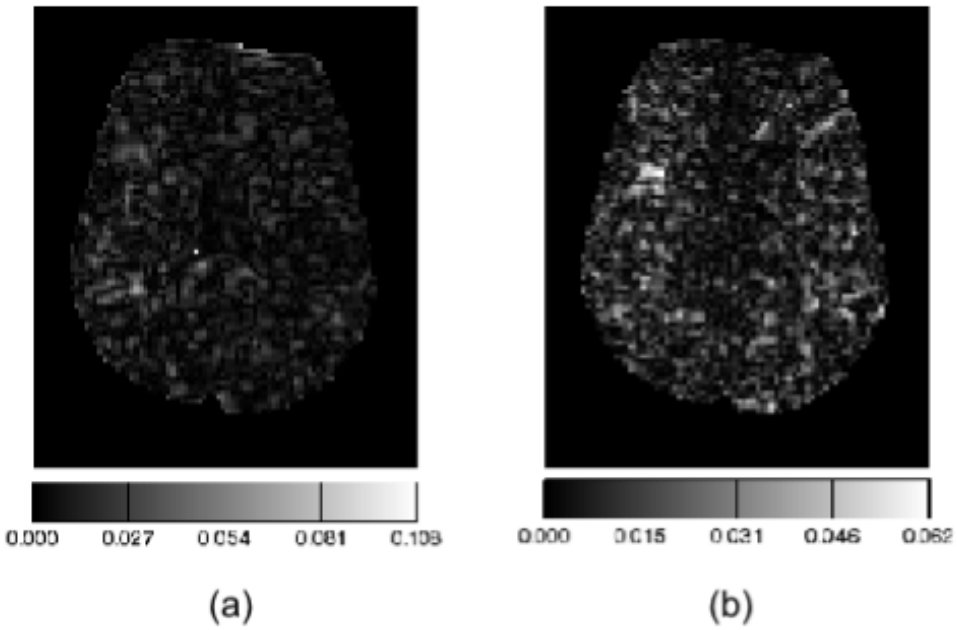


Figure 7.7: Part (a): relative error (absolute value of the difference divided by the correct value) between the trace of the diffusion tensor computed with and without the variance estimate given by formula (7.8). Part (b): absolute value of the difference between fractional anisotropy values computed with and without the variance estimated by formula (7.8).

pattern if a single value for the image intensity variance is used during tensor estimation. The pattern can be explained by the non-uniform intensity variance introduced by the image interpolation step. The patterns disappear when the correct noise variance in each voxel of each image, given by equation (E8), is used to compute the diffusion tensor. The striped pattern in the $\chi^2$ values is negligible if the DWI dataset was not aligned to a standard template, in addition to being corrected for motion and distortion, even if a single value for the intensity variance is used in estimating the tensor model. Nonetheless, formula (7.8) should be used in this case–because the images have suffered interpolation– to ensure an estimation of the correct variance values. Our results showed that in general the $\chi^2$ computed from registered images is lower than the $\chi^2$ computed from unregistered images when significant misregistration due to motion was present. However, the $\chi^2$ values computed using a single variance value estimated from the original (unregistered) images were lower than the $\chi^2$ values computed using equation (7.8) to estimate the correct intensity variance. This is to be expected since the variance of registered images at any given voxel location is less than or equal to the variance of the original (unregistered) images because of the interpolations necessary for registration. Thus, if a single variance value estimated from original (unregistered) images is used for the tensor computation, the overall effect will be an artificial decrease in the $\chi^2$ maps derived from the tensor fitting.

We have also shown that the estimation of the trace and fractional anisotropy parameters of the diffusion tensor can be affected by incorrect noise variance estimates. In the experiment shown, the error between the parameters estimated with and without the variance correction to account for image registration was

small: a few relative percentage points for the trace of the diffusion tensor and a few absolute percentage points for the fractional anisotropy index. We expect that the error caused by inappropriate weights in computing the actual parameters of the diffusion tensor model will be largest when the data being fit differs substantially from the model being used. To understand this, one only has to think of the extreme case in which the model fits the data without error. In this case, the weights being used become irrelevant since the numerator of the chi-squared equation becomes zero. The error between the data and the model arises from normally distributed thermal noise, physiological noise, as well as regions where it is known that the DT model poorly describes the underlying diffusion process, e.g. regions of crossing fibers. When considering only thermal, normally distributed additive noise, as we do throughout this paper, errors caused by incorrect variance estimates are not expected to be large and may diminish as the number of diffusion weighted images increases. As shown in the results section, however, these errors are expected to be in the order of a few percent.

The precise effect that changed image noise properties due to interpolation or approximation will have on DT estimation procedures cannot be determined a priori and will depend on several aspects of the registration and data processing procedures. Some of these are: the spatial transformations used to register the images, the interpolation or approximation kernel used, the noise variance and covariance in the original images, and the anatomical content of the images. However, it is worth noting that a translation of 0.5 pixels in all three dimensions can cause the variance of the signal to be reduced to 0.125 of the original variance of the signal when the linear interpolation method is used and if the data are spatially uncorrelated. If the correct noise variance value is not used, the resultant

$\chi^2$ measure will be underestimated by 8 times. Using the correlation matrix stated in equation (7.9) a translation of 0.5 pixels in all three dimensions would cause the variance in the interpolated image to be 0.25 of the variance in the original data. This would cause the $\chi^2$ measure to be underestimated by 4 times if all images in the dataset suffered similar interpolation.

Note that, although noise variance, and thus covariance, may vary even between datasets acquired using the same magnet and reconstructed using the same procedure, because of receive coil temperature or amplification settings for example, the noise correlation should not vary greatly. This is because the most significant correlations are introduced almost entirely by data-independent post-processing operations performed during magnitude image reconstruction. Thus, we expect that the method we propose to estimate the noise correlations in the images to be well suited when the MR images are reconstructed using the same procedure.

By inspecting images (a) and (c) displayed in figure 7.4 closely the reader may notice a slight vertical dark band running through the center of the images. We believe that this is due to non-uniformity in noise variance through the field of view caused by noise aliasing in the frequency encode direction (logical $x$ direction) during Fourier transform-based image reconstruction. The magnet receiver chain includes an analog filter, A/D converter, and a digital decimation filter. The filters reduce the response to higher frequencies. If the filters are not properly chosen, high frequency noise will be aliased into the Nyquist band. The observed pattern reflects the shape of the filter. There is no modulation of the brightness of an object in the field because the object fits into the FOV, so no aliasing takes place. There is no modulation of the noise in the phase encode

direction due to its low bandwidth, 100x lower than in the frequency-encode direction. We are in the process of determining the exact causes and remedies for the problems outlined above, but we do not believe that the slight (though noticeable) pattern in the noise materially affects our results.

### 7.5.1 Implications for analysis of variance of DT parameters

Knowledge of the uncertainty in the estimated diffusion tensor model parameters is important for assessing the significance of results of inter-subject or inter-acquisition comparisons. It is worth also noting that thermal noise variance not only plays a role in estimating the parameters of the model but also their uncertainty. From [17] it is known that when multivariate log-linear regression is used to compute the diffusion tensor parameters, the error variances of the estimated diffusion parameters are given by the diagonal elements of the matrix $\left(\mathbf{B}^T\Sigma_{\mathbf{e}}^{-1}\mathbf{B}\right)^{-1}$ (see appendix B), where $\mathbf{B}$ is the design matrix for the experiment, computed from the vectors that define the diffusion weighting gradients being used, and the diagonal values of $\Sigma_{\mathbf{e}}^{-1}$ given by $\tilde{S}_k^2/\mathrm{Var}(\tilde{S}_k)$, where $\tilde{S}_k$ represents the intensity value of the $k^{\mathrm{th}}$ image (for a fixed spatial coordinate) in the experiment. As shown in appendix B, if incorrect values of $\mathrm{Var}(\tilde{S}_k)$ are used the variance of the estimated parameters is no longer $\left(\mathbf{B}^T\Sigma_{\mathbf{e}}^{-1}\mathbf{B}\right)^{-1}$ and is given by equation (7.26). Methods for estimating the uncertainty in parameters computed through nonlinear models usually rely on Monte Carlo-type simulations for which it is necessary to know the variance that characterizes the uncertainty of each image intensity value [21].

### 7.5.2 Implications for DT-MRI-based tractography

One application which may be particularly affected by incorrect estimates intensity variance due to random thermal noise is DT-MRI-based tractography. Intensity variations due to thermal noise cause uncertainty in the orientation of greatest diffusivity measured in a DTI experiment. This uncertainty is normally computed using bootstrap [59] or Monte Carlo methods [21]. Such approaches are general in the sense that they can be used with both linear and nonlinear regression methods. On the down side, they are computationally intensive. In addition, bootstrap methods such as the one discussed in [59] require the acquisition of an additional amount of data. In both cases, testing the effect of different experimental setups (diffusion weighted directions, diffusion weighting strength, number of image replicates, etc.) can be cumbersome.

Alternatively, given a specific set of B-matrices and a diffusion tensor, a root mean square estimate of the uncertainty in orientation as a function of thermal noise variance can be derived using the theory of linear regression (see appendix B). This result can be used to calculate the approximate effect that incorrect intensity variance estimates can have on the variability of the principal diffusivity direction. We use a set of 22 b-matrices derived using the scheme described in [60], and an anisotropic diffusion tensor specified by the eigenvectors $\mathbf{g}_1 = \{1,0,0\}, \mathbf{g}_2 = \{0,1,0\}, \mathbf{g}_3 = \{0,0,1\}$, and eigenvalues $d_1 = 1,685 \times 10^{-6}, d_2 = 287 \times 10^{-6}, d_3 = 109 \times 10^{-6} mm^2/s$, and SNR = 15 to demonstrate the following example. When correct variance values are used in the estimation process the covariance matrix of the estimated DT parameters is given by equation (7.27) and the root mean square estimate in angle deviation when using the correct variance values is about 2.5 degrees. If the entire set of im-

ages is translated by 0.5 pixels in all three dimensions, using linear interpolation and using the covariance matrix stated in equation (7.9) for example, and the variance of each intensity value is not recomputed using the method described above, the covariance matrix of the estimated parameters is given by equation (7.26). The root mean square estimate of angle deviation in this case increases to about 10 degrees. This result seems counter intuitive since data interpolation should reduce the intensity variance of image values. This, in turn, should reduce the variability of the measurement of principal direction. We point out, however, that this is only caused by neglecting to account for the variance reduction due to image interpolation. If the variances of the image intensity values are appropriately re-calculated, the root mean square of the angle variation is reduced to about 1.3 degrees. We point out that these results are only approximations since they were obtained using first order expansion methods. Moreover, only variability due to thermal noise was included. However, it seems clear that tractography methods that rely on information about the variability of diffusivity orientation should be directly and adversely affected by neglecting to recomputed intensity variance estimates after image registration.

Moreover, we point out that probabilistic tractography is not the only approach that could be affected by intensity variance modifications due to interpolation. The deterministic methods presented in [27, 18] rely on estimating a continuous version of the diffusion tensor field for numerically computing the continuous path of presumed fiber tracts. In the approach described by Basser *et al.* [18] the continuous tensor field is estimated using an approximate fit to the discretely sampled diffusion tensor data using cubic b-splines. Conturo *et al.* [27] obtain a continuous version of the diffusion tensor field by interpolating

the diffusion weighted images where needed and fitting the DT model using the interpolated values. In both cases, the continuous approximation of diffusion tensor data produces diffusion tensors with different variance properties at different locations in the domain of the images. The variance of the interpolated diffusion tensors can be computed using the formulas given in appendix A. Thus such deterministic tract following approaches effectively integrate tracts by using estimated principal diffusivity directions that have different orientation uncertainty across different parts of the images, whether or not the raw DW images used for computing each diffusion tensor have been registered. At this point it is unclear what are the effects of non-uniform variance for such deterministic tract following methods. However, we believe that further investigation in the area is merited.

### 7.5.3 Implications for functional MRI and voxel based morphometry

Note that though we used diffusion tensor imaging as a case study, we believe that the same methodology could be used whenever data analysis requiring noise variance estimates is performed on registered or interpolated data. Some application examples in biomedical imaging include fMRI data analysis, studies of tissue shape and composition using statistical analysis of image data, MR relaxometry experiments, etc. In all such applications the goal is to detect image intensity changes that are the result of some biologically relevant phenomena. In fMRI this may be BOLD activation correlated with some type of brain activity, while in voxel based morphometry, for example, this may be information related to diseased tissue. Both fMRI data analysis and voxel based morphometry methods often rely on a generalized linear model for identifying the presence, absence,

and quantification, of biologically relevant phenomena. In this framework the measured image data (at a fixed voxel coordinate), defined by an $N$ dimensional vector $\mathbf{y}$, is modeled as a linear combination of explanatory coefficients arranged in an $N$x$M$ matrix $\mathbf{M}$ and unknown parameters defined by an $M$ dimensional vector $\mathbf{a}$: $\mathbf{y} = \mathbf{Ma} + \mathbf{e}$, where $\mathbf{e}$ represents an $N$ dimensional error vector whose entries are usually assumed to be independent, equally and normally distributed. If the error values are indeed normally distributed the maximum likelihood estimate for the model parameters is given by $\mathbf{a} = \left(\mathbf{M}^T\mathbf{M}\right)^{-1}\mathbf{M}^T\mathbf{y}$, while the covariance matrix of the estimates is given by $\mathbf{S_a} = \mathbf{LS_yL}^T$, with $\mathbf{L} = \left(\mathbf{M}^T\mathbf{M}\right)^{-1}\mathbf{M}^T$ and $\mathbf{S_a}, \mathbf{S_y}$ representing the covariance matrix of the estimated parameters and original data, respectively. Since the measurements $\mathbf{y}$ are usually assumed to be independently and identically distributed, the covariance matrix of the estimated parameters reduces to $\mathbf{S_a} = \lambda^2\left(\mathbf{M}^T\mathbf{M}\right)^{-1}$, with $\lambda^2$ being the assumed noise variance. Note that this analysis is usually performed on registered images in order to account for patient motion and geometric distortions. As shown in this paper, since different images will have different spatial transformations (and thus different interpolation) applied on them, the constant noise variance assumption is no longer appropriate. That is, the variance due to noise of an image value that has suffered interpolation is expected to be different from the variance of an image value that has suffered no interpolation at all. At this point it is unclear what effect this will have on image analysis results obtained using the general linear model, though it is an issue that should be investigated further.

## 7.6 Summary and conclusions

As fitting and estimating procedures from registered image data become increasingly more elaborate and quantitative, knowledge of the intensity variance due to noise will become more important for increasing the accuracy and scientific value of the results obtained from them. A method for estimating the variance in registered images is presented. The general approach can be summarized as follows. The output of the registration procedure is computed using an image interpolation or approximation procedure. The interpolation or approximation procedure can be written as a linear combination of the values of the image being registered. The coefficients of the linear combination are determined by the choice of interpolation or approximation kernel. Since the values of the image being registered are typically corrupted by noise, this operation can be viewed as a linear combination of random variables. The variance of the linear combination is given by well known statistical formulas.

The image interpolation or approximation generally required by image registration procedures will inevitably affect the noise variance properties of the images. We have shown that incorrect variance estimates can have a significant effect on diffusion tensor estimation procedures. The method we proposed for estimating the noise variance in registered images was shown to be successful in both simulated and real data experiments. Since $\chi^2$ measures and noise variance estimates are used more and more frequently in diffusion data analysis–examples include image registration [8], diffusion model selection [5, 99], robust tensor estimation [25], and brain tumor pathology detection [71]–correct variance estimates from registered image data will become increasingly important.

The methods described here could also be useful in other biomedical imaging applications such as MR relaxometry, fMRI data analysis, voxel based morphom-

etry, etc. However, the effects of the technique in each of these applications are not discussed in detail here and could be the subject of future study. The techniques described here could also find applications in other image processing and data analysis fields such as automatic target recognition and segmentation of registered data obtained from satellite or other remote sensing machinery. Statistical approaches are often used to fuse information gathered from several sensors and extract possible target matches.

## 7.7 Appendix A

We expect that different interpolation or approximation kernels will modify the variance in the registered images differently. The precise manner in which the choice of interpolator will affect the variance of an image is currently being investigated [90]. Here we give a general formula for the variance of the image intensity value produced using any kernel-based interpolation method due to a spatial transformation being applied during registration. Let $s(\mathbf{k})$, where $\mathbf{k} \in \mathbb{Z}^d$, with $d$ being the dimension of the images, be the discretely sampled image produced by the acquisition system. A continuous approximation to $s(\mathbf{k})$ is given by:

$$\tilde{s}(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{Z}^d} s(\mathbf{k}) h(\mathbf{x} - \mathbf{k}) \tag{7.12}$$

Note that the summations are carried from $-\infty$ to $\infty$ by making the images periodic. Naturally, if we would like the values $\tilde{s}(\mathbf{x})$ to be equal to the values of $s(\mathbf{k})$ at coordinates $\mathbf{x} = \mathbf{k}$, then $h(\mathbf{x})$ must have the following properties:

$$h(\mathbf{k}) = 0 \ \forall \mathbf{k} \neq 0 \tag{7.13}$$

and

$$h(0) = 1. \tag{7.14}$$

Examples of such kernels are the linear hat function (also known as B-spline of order 1), and the popular sinc kernel given by:

$$h(x) = \begin{cases} 1 - |x| & \text{if } |x| \leq 1; \\ 0 & \text{if } |x| > 1. \end{cases} \tag{7.15}$$

and

$$h(x) = \frac{\sin(\pi x)}{\pi x} \tag{7.16}$$

respectively. Note that true sinc interpolation is almost never used in the field of medical imaging because of the enormous computational cost associated with it. Since the support of sinc is infinite, in theory, the sum in (7.12) should be evaluated from $-\infty$ to $\infty$. Because of such computational costs and other reasons (i.e. ringing artifacts) researchers in the field prefer to use truncated and apodized versions of (7.16) [46, 66, 108, 75]. Note also that in the cases where $d > 1$, the interpolation kernel is replaced by

$$\hat{h}(\mathbf{x}) = \prod_{i=1}^{d} h(x_i). \tag{7.17}$$

If the basis function being used does not satisfy the properties stated in equations (7.13) and (7.14), examples include the popular B-splines of order 2 or greater, equation (7.12) needs to be adjusted. Let $b(x)$ be a basis function such that properties (7.13) and (7.14) do not hold. The interpolation equation then becomes:

$$\tilde{s}(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{Z}^d} c(\mathbf{k}) b(\mathbf{x} - \mathbf{k}). \tag{7.18}$$

The coefficients $c(\mathbf{k})$ are given by:

$$c(\mathbf{k}) = \left(b^{-1} * s\right)(\mathbf{k}), \tag{7.19}$$

where $b^{-1}$ is the uniquely defined convolution-inverse [114, 115]. As shown in [114, 115], we can substitute (7.19) into (7.18) to see that

$$\begin{aligned}\tilde{s}(\mathbf{x}) &= \sum_{\mathbf{k}\in\mathbb{Z}^d} \left(b^{-1} * s\right)(\mathbf{k}) b(\mathbf{x}-\mathbf{k}) \\ &= \sum_{\mathbf{k}_1\in\mathbb{Z}^d} \sum_{\mathbf{k}_2\in\mathbb{Z}^d} b^{-1}(\mathbf{k}_2) s(\mathbf{k}_1-\mathbf{k}_2) b(\mathbf{x}-\mathbf{k}_1) \\ &= \sum_{\mathbf{k}\in\mathbb{Z}^d} s(\mathbf{k}) h(\mathbf{x}-\mathbf{k})\end{aligned} \tag{7.20}$$

where the new interpolation kernel is given by:

$$h(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} b^{-1}(\mathbf{k}) b(\mathbf{x}-\mathbf{k}). \tag{7.21}$$

Thus, the variance of the interpolated image intensity value due to spatial transformation $f(\mathbf{x})$ is given by:

$$\begin{aligned}\mathrm{Var}\left(S(f(\mathbf{x}))\right) &= \sum_{\mathbf{k}\in\mathbb{Z}^d} \mathrm{Var}\left(S(\mathbf{k})\right) \left(h(f(\mathbf{x})-\mathbf{k})\right)^2 + \\ &\quad \sum_{\mathbf{i}\in\mathbb{Z}^d} \sum_{\mathbf{j}\in\mathbb{Z}^d, \mathbf{j}\neq\mathbf{i}} h(f(\mathbf{x})-\mathbf{j}) h(f(\mathbf{x})-\mathbf{i}) \mathrm{Cov}\left(s(\mathbf{i}), s(\mathbf{j})\right)\end{aligned} \tag{7.22}$$

Applications such as geometrical distortion correction due to imperfect magnetic fields in MRI require the formula above to be multiplied by the square of the determinant of the Jacobian matrix of $f$, as in equation (7.8):

$$\begin{aligned}\mathrm{Var}\left(S(f(\mathbf{x}))\right) &= \gamma \sum_{\mathbf{k}\in\mathbb{Z}^d} \mathrm{Var}\left(S(\mathbf{k})\right) \left(h(f(\mathbf{x})-\mathbf{k})\right)^2 + \\ &\quad \gamma \sum_{\mathbf{i}\in\mathbb{Z}^d} \sum_{\mathbf{j}\in\mathbb{Z}^d, \mathbf{j}\neq\mathbf{i}} h(f(\mathbf{x})-\mathbf{j}) h(f(\mathbf{x})-\mathbf{i}) \mathrm{Cov}\left(s(\mathbf{i}), s(\mathbf{j})\right).\end{aligned} \tag{7.23}$$

with $\gamma = \left(\det\left|Jac(f(\mathbf{x}))\right|\right)^2$.

## 7.8 Appendix B

Using the log-linear diffusion tensor model for the diffusion weighted image data we analyze the error distribution of the estimated diffusion tensor parameters.

Let $\mathbf{y} = \{\ln(S_1), , \ln(S_N)\}^T$, where $S_i$ represents the ith measurement in a typical DTI acquisition, and $\mathbf{a} = \{D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}, \ln(A_0)\}^T$ represent the diffusion tensor model parameters. To first order, the log linear model can be written as

$$\mathbf{y} = \mathbf{B}\mathbf{a} + \mathbf{e} \tag{7.24}$$

where the $j^{\text{th}}$ row of $\mathbf{B}$ is composed of the b-matrix entries of the $j^{\text{th}}$ diffusion weighted acquisition parameters $\{-b_{xxj}, -b_{yyj}, -b_{zzj}, -2b_{xyj}, -2b_{xzj}, -2b_{yzj}, 1\}$, and $\mathbf{e}$ represents the error vector. The covariance matrix of $\mathbf{e}$ is denoted $(\Sigma_{\mathbf{e}})_{ii} = \sigma_i^2/\langle S_i \rangle^2$, where $\langle u \rangle$ denotes the expectation of random variable $u$. Since each measured data point in $\mathbf{y}$ was taken independently at different times $(\Sigma_{\mathbf{e}})_{ij} = 0 \; \forall i \neq j$. All terms in equation (7.24) are considered deterministic, with exception of $\mathbf{e}$ which represents error due to noise in the imaging acquisition system. Therefore $\Sigma_{\mathbf{y}} = \Sigma_{\mathbf{e}}$. In practice however, one can only estimate $\Sigma_{\mathbf{e}}$. This is usually done based on measurements from background intensity values. As shown in this paper, the estimates of the variance in each image intensity value need to be re-calculated after registration or interpolation is performed on the images. We will differentiate the true covariance matrix of the data $\Sigma_{\mathbf{e}}$ from the estimated one $\tilde{\Sigma}_{\mathbf{e}}$. The weighted least squares solution to equation (7.24) is given by:

$$\mathbf{a} = \left(\mathbf{B}^T \tilde{\Sigma}_{\mathbf{e}}^{-1} \mathbf{B}\right)^{-1} \left(\mathbf{B}^T \tilde{\Sigma}_{\mathbf{e}}^{-1}\right) \mathbf{y} \tag{7.25}$$

while the covariance matrix of the estimated parameters is given by:

$$\Sigma_{\mathbf{a}} = \left(\mathbf{B}^T \tilde{\Sigma}_{\mathbf{e}}^{-1} \mathbf{B}\right)^{-1} \mathbf{B}^T \tilde{\Sigma}_{\mathbf{e}}^{-1} \Sigma \mathbf{e} \tilde{\Sigma}_{\mathbf{e}}^{-1} \mathbf{B} \left(\mathbf{B}^T \tilde{\Sigma}_{\mathbf{e}}^{-1} \mathbf{B}\right)^{-1} \tag{7.26}$$

If our estimate of the covariance matrix of the measured data is precise and accurate, $\tilde{\Sigma}_{\mathbf{e}} \approx \Sigma_{\mathbf{e}}$, then equation (7.26) reduces to:

$$\Sigma_{\mathbf{a}} = \left(\mathbf{B}^T \tilde{\Sigma}_{\mathbf{e}}^{-1} \mathbf{B}\right)^{-1}. \tag{7.27}$$

If, on the other hand, errors are made in calculating $\tilde{\Sigma}_{\mathbf{e}}$, such as neglecting to account for the interpolation applied to the data during image registration, the covariance of the estimated parameters is given by equation (7.26).

The uncertainty in the principal direction orientation in a diffusion tensor $\tilde{D}$ calculated using (7.25) can be estimated by studying the effects of random perturbations $\triangle D$ on a deterministic tensor $D_0$ [15]:

$$\tilde{D} = D_0 + \triangle D. \tag{7.28}$$

Let $d_1, d_2, d_3$, and $\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3$, represent the eigenvalues (arranged in decreasing order) and eigenvectors, respectively, of the three dimensional positive definite symmetric tensor $D_0$. We are interested in computing the perturbation $\tilde{\mathbf{g}}_1 = \mathbf{g}_1 + \triangle \mathbf{g}_1$. We will assume that $D_0$ comes from biological tissue with high diffusion anisotropy so that sorting bias in the computed eigenvalues can be safely neglected. It can be shown [39] that, to first order, the perturbation of the eigenvector associated with greatest diffusivity is:

$$\triangle \mathbf{g}_1 = \sum_{i=2}^{3} \left( \frac{\mathbf{g}_1^T \triangle D \mathbf{g}_i}{d_1 - d_i} \right) \mathbf{g}_i. \tag{7.29}$$

The perturbation angle $\theta$ between $\tilde{\mathbf{g}}_1$ and $\mathbf{g}_1$ is thus $\theta = \tan^{-1}(\|\triangle \mathbf{g}_1\|)$ (Basser 1997). Noting that the eigenvectors $\mathbf{g}$ form an orthonormal basis for the 3D Euclidean space and using the small angle approximation for $\tan\theta$:

$$\theta \approx \| \triangle \mathbf{g}_1 \| = \sqrt{\sum_{i=2}^{3} \left( \frac{\mathbf{g}_1^T \triangle D \mathbf{g}_i}{d_1 - d_i} \right)^2}. \tag{7.30}$$

Now

$$\langle \theta^2 \rangle \approx \langle \| \triangle \mathbf{g}_1 \|^2 \rangle = \left\langle \sum_{i=2}^{3} \left( \frac{\mathbf{g}_1^T \triangle D \mathbf{g}_i}{d_1 - d_i} \right)^2 \right\rangle = \sum_{i=2}^{3} \frac{\left\langle (\mathbf{g}_1^T \triangle D \mathbf{g}_i)^2 \right\rangle}{(d_1 - d_i)^2} \tag{7.31}$$

while, as shown in [6],

$$\left\langle (\triangle D_{ii})^2 \right\rangle = \left\langle (\tilde{D}_{ii} - (D_0)_{ii})^2 \right\rangle = (\Sigma_{\mathbf{a}})_{i,i} \tag{7.32}$$

and similarly

$$\left\langle (\triangle D_{ij})^2 \right\rangle = \left\langle (\tilde{D}_{ij} - (D_0)_{ij})^2 \right\rangle = (\Sigma_{\mathbf{a}})_{i+j+1,i+j+1}. \tag{7.33}$$

Looking at the principal axis case, $\mathbf{g}_i^T \triangle D \mathbf{g}_j = \triangle D_{ij}$. In general, the covariance matrix can be rotated so that $\left\langle (\mathbf{g}_i^T \triangle D \mathbf{g}_j) \right\rangle = (\mathbf{R}\Sigma_{\mathbf{a}}\mathbf{R}^T)_{i+j+1,i+j+1} = \Xi_{i+j+1,i+j+1}$ [6], where $\mathbf{R}$ is a rotation matrix. Therefore (7.31) can be written as

$$\left\langle \|\triangle \mathbf{g}_1\|^2 \right\rangle = \sum_{i=2}^{3} \frac{\left\langle (\mathbf{g}_1^T \triangle D \mathbf{g}_i)^2 \right\rangle}{(d_1 - d_i)^2} = \sum_{i=2}^{3} \frac{(\Xi)_{2+i,2+i}}{(d_1 - d_i)^2}. \tag{7.34}$$

Thus the root mean square angle estimate of the deviation from the principal direction is:

$$\theta_{\mathrm{RMS}} = \sqrt{\sum_{i=2}^{3} \frac{(\Xi)_{2+i,2+i}}{(d_1 - d_i)^2}}, \tag{7.35}$$

where the covariance matrix of the estimated parameters $\Sigma_{\mathbf{a}}$ in $\mathbf{R}\Sigma_{\mathbf{a}}\mathbf{R}^T = \Xi$ is given in equation (7.26).

# Chapter 8

# Summary and Conclusions

Misregistration artifacts typically stem from patient (object) or camera motion, as well as device dependent geometric distortions. As shown with real experimental data in chapters 5 and 6, misregistration-related artifacts can pose significant difficulties to the automated measurement of quantitative parameters such as diffusion coefficients in diffusion weighted MRI experiments, or activation potential duration in epicardial optical imaging experiments. Such artifacts are not limited to the two imaging experiments mentioned in detail but also occur in functional MRI experiments, or any time data from images taken at different times or by different sensors are used to estimate some kind of physical quantity.

In typical quantitative imaging experiments a sequence of processing steps is employed to translate the measured data into the desired quantitative information. This sequence of steps can include a tomographic reconstruction step to transform the measured data into images, as well as several post processing steps employed to remove misalignment artifacts, in addition to other artifacts such as intensity inhomogeneities and system noise. Once the undesirable effects of motion, distortion, and noise in the image data have been been accounted for, the series of image data is then used in a data analysis step whose goal is to extract

interesting quantitative parameters usually via model fitting and other statistical methods. The extracted parameters are then used for a variety of important tasks. In the biomedical world, these results are used for medical diagnosis, surgical planning, treatment monitoring, image guided treatment, as well as basic science research. In robotics and military applications, the processed data can be used for automated tracking, target recognition and classification, as well as surveillance purposes.

Presented in this work were several novel ideas related to the measurement and removal of undesirable misalignment artifacts in typical imaging experiments. The novel contributions presented throughout this dissertation were the fruit of an overall philosophy taken towards the problem at hand: that in order to achieve a specific measurement obtained from an imaging experiment it is best to consider explicitly all steps taken in a typical image processing pipeline. This is due to the fact that the output from any given processing procedure can have severe impact on the next step in the processing pipeline. Even though a great amount of research on image processing methods for several aspects of typical imaging experiments has been performed over the past two or so decades (over 180 journal publications solely devoted to biomedical image registration were published just in 2004) the majority of these works, however, give minimal consideration to the relation between the particular methodology being explored and other processing steps in a typical image processing pipeline. Too often researchers are willing to accept important steps such as image acquisition, reconstruction, denoising, etc., as 'black boxes' in the process of developing the latest registration or segmentation algorithm. This is understandable since quantitative imaging efforts typically involve scientists with a variety of backgrounds. On the data acquisition side

there are imaging physicists and engineers who are concerned with many technical issues relating to the physics of the data measurement process. Their goal is to try to expand the capability of present systems by improving sampling rates as well as minimizing system noise and other artifacts. At the other end of the spectrum (in the biomedical fields) there are neuroscientists and physicians who wish to use the results from quantitative experiments to better understand biological phenomena, maximize treatment success, etc. In between there are data processors whose background include a variety of fields such as applied mathematics, statistics, electrical engineering, as well as scientific computation. Their goal is try to remove artifacts which cannot be controlled at the time of data acquisition and to design and conduct computations that will provide the desired quantitative outputs. Given their varied background and expertise, communication between researchers at different stages of the imaging experiment can be less than optimal. This, combined with the ever increasing urge to publish results as quickly as possible, can cause many disconnects which may delay progress in many fields of experimental sciences. As we saw in chapters 4 and 7, the common disregard for the stochastic nature of the imaging acquisition process, for example, caused researchers to perform sub-optimal computations for many years.

The specific scientific contributions presented in this work included both analytical and computational advances and can be summarized as follows. A linear, shift-invariant, stochastic image model was used to better understand the limitations of current methodology used for registering multiple images as well as for computation of quantitative parameters of interest. Methods for improving on such limitations were also described. In chapter 4 we described the effect of system noise on the computations necessary for registering multiple images. It

was shown that system noise will inevitably introduce systematic local optima artifacts in intensity-based automatic registration methods. Such artifacts will prevent optimization methods from finding the global optimum for any particular problem. In addition, even when approximate locations for the global optimum are known, the computation methods normally used will still give biased answers: the solution to the registration problem will be biased towards the most 'blurred' image. This is due to the fact that common energy functionals used for guiding the registration process are sensitive to the stochastic properties of the input images. These, in turn, are sensitive to the spatial transformations being applied as well as the continuous image model chosen. Solutions for mitigating such artifacts, such as prior data 'blurring' and the use of sinc-based basis functions, were presented and discussed.

Next, in chapter 5, a comprehensive solution to the motion and distortion artifacts encountered when processing diffusion weighted MRI of the human brain was presented. The computational solution employed the data 'blurring' strategy to minimize local optima artifacts. In addition, the linear shift-invariant imaging model was used to derive the effects of both patient motion and eddy current-induced distortions explicitly so that computation of parameters of interest (in this case the diffusion tensor) can be done properly. The approach was tested using real data from various experiments and it was found that it is capable of removing even severe motion and distortion related artifacts. A similar technique was used in chapter 6 to remove artifacts related to cardiac motion in 2D epicardial fluorescence optical imaging experiments. Because of the effects motion and distortion in linear, shift-invariant, imaging systems described in chapter 3, intensity corrections were not used even though cardiac tissue motion is tech-

nically non-rigid. Results showed that the techniques used are not limited to MRI-type instruments but can also be applied in cases where data comes from optical sensors.

Finally, in chapter 7 methods for computing quantitative parameters of interest once the transformations that align the images are known were presented. Until now the effects of image resampling and interpolation on the stochastic properties of the images, and how these affect popular data analysis methods such as the maximum likelihood method, were missing. These were discussed in chapter 7 and it was concluded that neglecting the effects of the data processing can cause errors of large magnitude in the computation of the quantities of interest, as well as their associated uncertainty.

By investigating the effects of specific processing steps such as image registration on simple models for the image formation process one can better understand the limitations of current image processing methodology. Naturally, conclusions have to be limited to situations in which the image formation model chosen is deemed appropriate. For example, it is not clear that the work presented here will be valid for other types of system noise such as multiplicative noise, or for nonlinear image formation processes. It should also be understood that none of the work presented herein validates, in any strict sense, the image formation models used to explain the different phenomena in the specific imaging modalities mentioned (MRI and CCD). At most it can be concluded that some of the experimental results obtained are consistent with some of the aspects derived using the linear stochastic image formation model.

However, the idea of using specific image formation models to understand the limitations of several processing steps in an imaging experiment seems to

be a powerful alternative to many experimental studies whose goal is to make generalizations based on empirical evidence. For example, several researchers had attempted to understand the so called 'grid' artifacts in mutual information-based registration curves [70, 85, 112, 117]. These studies, to a large extent, were based on the attempt to investigate the effects of different images and image interpolation methods, experimentally, on mutual information registration curves. While some knowledge can be obtained from such studies, the extent to which conclusions from them obtained hold will be unclear. It is hoped that the use of image models to investigate the performance of image processing techniques such as registration, denoising, and segmentation, to name a few, will become an increasingly popular trend. More generally, researchers in quantitative imaging fields may benefit from viewing tasks related to artifact removal and parameter estimation within the context of 'image science' rather than 'image processing.'

# BIBLIOGRAPHY


[1] US Patent number 4,698,591: Method for magnetic field gradient eddy current compensation, 1987.

[2] US Patent number 4,737,716: Self-shielded gradient coils for nuclear magnetic resonance imaging, 1988.

[3] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover, 1974.

[4] A. Aldroubi, M. Unser, and M. Eden. Cardinal spline filters: Stability and convergence to the ideal sinc interpolator. *Signal Processing*, 28:127–138, 1992.

[5] D.C. Alexander, G.J. Barker, and S.R. Arridge. Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48:331–340, 2002.

[6] A.W. Anderson. Theoretical analysis of the effects of noise on diffusion tensor imaging. *Magnetic Resonance in Medicine*, 46:1174–1188, 2001.

[7] T.W. Anderson. *An Introduction to Multivariate Statistical Analysis*, chapter Principal Components. John Wiley and Sons, New York, 1957.

[8] J.L.R. Andersson and S. Skare. A model-based method for retrospective correction of geometric distortions in diffusion-weighted EPI. *Neuroimage*, 16:177–199, 2002.

[9] J. Ashburner and K.J. Friston. Image Registration. In C.T.W. Moonen and P.A. Bandettini, editors, *Functional MRI*. Springer-Verlag, 1999.

[10] J. Ashburner and K.J. Friston. Nonlinear spatial normalization using basis functions. *Humand Brain Mapping*, 7:254–266, 1999.

[11] J. Ashburner and K.J. Friston. Voxel-based morphometry–the methods. *Neuroimage*, 11:805–821, 2000.

[12] K. Astrom and A. Heyden. Stochastic analysis of image acquisition, interpolation and scale-space smoothing. *Adv. in Appl. Probab.*, 31:855–894, 1999.

[13] H.H. Barret and K.J. Myers. *Foundations of Image Science*. Wiley, Hoboken, New Jersey, 2004.

[14] P.J. Basser. Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR in Biomedicine*, 8:333–344, 1995.

[15] P.J. Basser. Quantifying errors in fiber-tract direction and diffusion tensor field maps resulting from mr noise. In *Proceedings of the ISMRM*, 1997.

[16] P.J. Basser and D.K. Jones. Diffusion tensor MRI: theory, experimental design and data analysis - a technical review. *NMR in Biomedicine*, 15:456–467, 2002.

[17] P.J. Basser, J. Mattielo, and D. LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66:259–267, 1994.

[18] P.J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, 44:625–632, 2000.

[19] P.J. Basser and C. Pierpaoli. Microstructural and physiological features of tissues elucidated by quantitative diffusion tensor MRI. *Journal of Magnetic Resonance B*, 111:209–219, 1996.

[20] M.E. Bastin. Correction of eddy current-induced artifacts in diffusion tensor imaging using iterative cross-correlation. *Magnetic Resonance in Medicine*, 17:1011–1024, 1999.

[21] T.E. Behrens, M.W. Woolrich, M. Jenkinson, H. Johansen-Berg, R.G. Nunes, S. Clare, P.M. Brady, and S.M. Smith. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance Medicine*, 50:1077–1088, 2003.

[22] F.L. Bookstein. Principal warps: thin-plate splines and the decomposition of transformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:567–585, 1989.

[23] F.L. Bookstein. Voxel-based morphometry should not be used with imperfectly registered images. *Neuroimage*, 14:1454–1462, 2001.

[24] BrainWeb. http://www.bic.mni.mcgill.ca/brainweb/.

[25] L.-C. Chang, G.K. Rohde, D.K. Jones, P.J. Basser, and C. Pierpaoli. Restore: Robust estimation of tensors by outlier rejection. In *Proceedings of the ISMRM*, 2004.

[26] G.E. Christensen, R.D. Rabbitt, and M.I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5:1435–1447, 1996.

[27] T.E. Conturo, N.F. Lori, T.S. Cull, E. Akbudak, A.Z. Snyder, J.S. Shimony, R.C. McKinstry, H. Burton, and M.E. Raichle. Tracking neuronal fiber pathways in the living human brain. *Proc. of the National Academies of Sciences USA*, 96:10422–10427, 1999.

[28] W.R. Crum, L.D. Griffin, D.L.G. Hill, and D.J. Hawkes. Zen and the art of medical image registration: correspondence, homology, and quality. *Neuroimage*, 20:1425–1437, 2003.

[29] W.R. Crum, T. Hartkens, and D.K.G. Hill. Non-rigid image registration: theory and practice. *The British Journal of Radiology*, 77:140–153, 2004.

[30] C. Davatzikos and R. Bryan. Using a deformable suface model to obtain a shape representation of the cortex. *IEEE Transactions on Medical Imaging*, 15:786–795, 1996.

[31] W.F. Eddy, M. Fitzgerald, and D.C. Noll. Improved image registration by using fourier interpolation. *Magnetic Resonance in Medicine*, 36:923–931, 1996.

[32] I.R. Efimov, D.T. Huang, J.M. Rendt, and G. Salama. Optical mapping of repolarization and refractoriness from intact hearts. *Circulation*, 90:1469=1479, 1994.

[33] A.C. Evans, C. Beil, S. Marrett, C.J. Thompson, and A. Hakim. Anatomical-functional correlation using an adjustable mri-based region of interest atlas with positron emission tomography. *Journal of Cerebral Blood Flow and Metabolism*, 8:512–530, 1988.

[34] A.C. Evans, S. Marrett, J. Torrescorzo, S. Ku, and L. Collins. Mri-pet correlation in three dimensions using a volume of interest (VOI) atlas. *J. Cerebral Blood Flow Metabolism*, 11:A69–A78, 1991.

[35] O. Faugeras and G. Hermosillo. Well-posedness of two nonrigid multimodal image registration methods. *SIAM J. Appl. Math.*, 64:1550–1587, 2004.

[36] L.R. Frank. Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 47:1083–1099, 2002.

[37] K.J. Friston, A.P. Holmes, K.J. Worsley, J.P Pioline, C.D Frith, and R.S. Frackowiak. Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2:189–210, 1995.

[38] K.J. Friston, S. Williams, R. Howard, R.S. Frackowiak, and R. Turner. Movement-related effects in fmri time-series. *Magnetic Resonance in Medicine*, 35:346–355, 1996.

[39] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, Inc., New York, 1972.

[40] Goldstein. *Classical Mechanics*. Addison-Wesley, 1950.

[41] R.A. Gray, A.M. Pertsov, and J. Jalife. Spatial and temporal organization during cardiac fibrillation. *Nature*, 392:75–78, 1998.

[42] S. Grootoonk, C. Hutton, J. Ashburner, A.M. 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*, 11:49–57, 2000.

[43] H. Gudbjartsson and S. Patz. The rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, 34:910–914, 1995.

[44] E.M. Haacke, R.W. Brown, M.R. Thompson, and R. Venkatesan. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Wiley-Liss, New York, 1999.

[45] J. V. Hajnal, D.L.G. Hill, and D.J. Hawkes. *Medical Image Registration*. CRC Press, New York, 2001.

[46] 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 ComputerAssisted Tomography*, 19:289–296, 1995.

[47] J.C. Haselgrove and J.R. Moore. Correction for distortion of echo-planar images used to calculate the apparent diffusion coefficient. *Magnetic Resonance in Medicine*, 36:960–964, 1996.

[48] Z. He, J.C. Maublant, J.C. Cauvin, and A. Veyre. Reorientation of the left ventricular long-axis on myocardial transaxial tomograms by a linear fitting method. *J. Nucl. Med.*, 32:1794–1800, 1991.

[49] R.M. Henkelman. Measurement of signal intensities in the presence of noise in MR images. *Medical Physics*, 13:232–233, 1985.

[50] G. Hermosillo, C. Chefd'Hotel, and O. Faugeras. A variational approach to multimodal image matching. *International Journal of Computer Vision*, 50:329–343, 2002.

[51] F.B. Hildebrand. *Methods of Applied Mathematics*. Dover Publications, 1992.

[52] D.L.G. Hill, D.J. Hawkes, J.E. Crossman, M.J. Gleeson, T.C.S. Cox, E.E.C.M.L. Bracey, A.J. Strong, and P. Graves. Registration of MR and CT images for skull base surgery using point-like anatomical features. *British Journal of Radiology*, 64:1030–1035, 1991.

[53] R.V. Hogg and A.T. Craig. *Introduction to Mathematical Statistics*. Prentice-Hall,Inc., Upper Saddle River, NJ., 1995.

[54] B.K.P. Horn and B.G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[55] M.A. Horsfield. Mapping eddy current induced fields for the correction of diffusion-weighted echo planar images. *Magnetic Resonance Imaging*, 17:1335–1345, 1999.

[56] W.M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1:35–51, 1996.

[57] B.A. Inglis, E.L. Bossart, D.L. Bucley, E.D. Wirth III, and T.H. Mareci. Visualization of neural tissue water compartments using biexponential diffusion tensor MRI. *Magnetic Resonance in Medicine*, 45:580–587, 2001.

[58] P. Jezzard, A.S. Barnett, and C. Pierpaoli. Characterization of and correction for eddy current artifacts in echo planar diffusion imaging. *Magnetic Resonance in Medicine*, 39:801–812, 1998.

[59] D.K. Jones. Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI. *Magnetic Resonance in Medicine*, 49:7–12, 2003.

[60] D.K. Jones, M.A. Horsfield, and A. Simmons. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magnetic Resonance in Medicine*, 42:515–125, 1999.

[61] S. Kullback. *Information Theory and Statistics*. Dover Publications, 1968.

[62] S. Kullback. *Information Theory and Statistics*, chapter Multivariate Normal Populations. Dover Publications, 1968.

[63] K.F. Kwaku and S.M. Dillon. Shock-induced depolarization of refractory myocardium prevents wave-front propagation in defibrillation. *Circ. Res.*, 79:957–973, 1996.

[64] 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.

[65] J. Kybic and M. Unser. Fast parametric elastic image registration. *IEEE Transactions on Image Processing*, 12(11):1427–1442, November 2003.

[66] T.M. Lehmann, C. Gonner, and K. Spitzer. Survey: Interpolation methods in medical image processing. *IEEE Transactions on Medical Imaging*, 18:1049–1075, 1999.

[67] S. F. Lin, B. J. Roth, and Jr. J. P. Wikswo. Quatrefoil reentry in myocardium: an optical imaging study of the induction mechanism. *Journal of Cardiovascular Electrophysiology*, 10:574–586, 1999.

[68] L.D. Lunsford. *Modern Stereotactic Neurosurgery*. Martinus Nijhoff, Boston, MA, 1988.

[69] L.C. Maas and P.F. Renshaw. Post-registration spatial filtering to reduce noise in functional MRI data sets. *Magnetic Resonance Imaging*, 17:1371–1382, 1999.

[70] F. Maes, A. Collignon, D. Vandermeulen, G. Marcha, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16:187–198, 1997.

[71] S.E. Maier, H. Mamata, and R.V. Mulkern. Characterization of normal brain and brain tumor pathology by chisquares parameter maps of diffusion-weighted image data. *European Journal of Radiology*, 45:199–207, 2003.

[72] J.B. Maintz and M.A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2:1–36, 1998.

[73] J.F. Mangin, C. Poupon, C. Clark, D. Le Bihan, and I. Block. Distortion correction and robust tensor estimation for MR diffusion imaging. *Medical Image Analysis*, 6:191–198, 2002.

[74] D. Meier and E. Fisher. Parameter space warping: shape-based correspondence between morphologically different objects. *IEEE Transactions on Medical Imaging*, 21:31–47, 2002.

[75] E.H.W. Meijering, W.J. Niessen, and M.A.Viergever. Quantitative evaluation of convolution-based methods for medical image interpolation. *Medical Image Analysis*, 5:111–126, 2001.

[76] J. Modersitzki. *Numerical Methods for Image Registration*. Oxford University Press, Oxford, 2004.

[77] S.G. Nash and A. Sofer. *Linear and Nonlinear Programming*. McGraw-Hill, New York, 1996.

[78] F. Natterer, M. Cheney, and B. Borden. Resolution for radar and x-ray tomography. *Inverse Problems*, 19:S55–S63, 2003.

[79] L.D. Nickerson, S. Narayana, J.L. Lancaster, P.T. Fox, and J.H. Gao. Estimation of the local statistical noise in positron emission tomography revisited: practical implementation. *Neuroimage*, 19:442–256, 2003.

[80] National Library of Medicine. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi.

[81] S. Pajevic and C. Pierpaoli. Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data. *Magnetic Resonance in Medicine*, 42:526–540, 1999.

[82] 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 brain. *Journal of Computer Assisted Tomography*, 13:20–26, 1989.

[83] C. Pierpaoli, P. Jezzard, P.J. Basser, A.S. Barnett, and G. Di Chiro. Diffusion tensor MR imaging of the human brain. *Radiology*, 201:637–648, 1996.

[84] J.P. Pluim, J. B. Mainz, and M.A. Vergever. Mutual-information-based registration of medical images: a surey. *IEEE Transactions on Medical Imaging*, 22:986–1004, 2003.

[85] 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.

[86] W.H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cabridge University Press, New York, 1988.

[87] G. K. Rohde, A. Aldroubi, and B. M. Dawant. The adaptive bases algorithm for intensity-based nonrigid image registration. *IEEE Transactions on Medical Imaging*, 22:1470–1479, 2003.

[88] G. K. Rohde, S. Pajevic, and C. Pierpaoli. Multi-channel registration of diffusion tensor images using directional information. Proceedings of the

2nd International Symposium on Biomedical Imaging, Arlington, VA, USA, 2004.

[89] G. K. Rohde, S. Pajevic, C. Pierpaoli, and P. J. Basser. A comprehensive approach for multi-channel image registration. Proceedings of 2nd International Workshop on Biomedical Image Registration, Philadelphia, PE, USA, pages 214–223, 2003.

[90] G.K. Rohde. Ph.D. dissertation research. Applied Mathematics and Scientific Computation Program. University of Maryland, College Park., 2004.

[91] G.K. Rohde, A.S. Barnett, P.J. Basser, S. Marenco, and C. Pierpaoli. Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 51:103–114, 2004.

[92] G.K. Rohde, A.S. Barnett, P.J. Basser, and C. Pierpaoli. Estimating intensity variance due to noise in registered images: Applications to diffusion tensor MRI. *Neuroimage*, 26:673–684, 2005.

[93] G.K. Rohde, B.M Dawant, and S.F. Lin. Correction of motion artifact in cardiac optical mapping using image registration. *IEEE Transactions on Biomedical Engineering*, 52:338–341, 2005.

[94] T. Rohlfing, C.R. Maurer Jr., W.G. O'Dell, and J. Zhong. Modeling liver motion and deformation during the respiratory cycle using intensity-based nonrigid registration of gated MR images. *Medical Physics*, 31:427–432, 2004.

[95] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes. Nonrigid registration using free-form deformations: application to

breast MR images. *IEEE Transactions on Medical Imaging*, 18:712–721, 1999.

[96] R. Schaback. *Creating surfaces from scattered data using radial basis functions, Mathematical Models for Curves and Surfaces*. Vanderbilt University Press, Nashville, 1995.

[97] P.H. Schönemann. A generalized solution of the orthogonal procrustes problem. *Psychometrica*, 31:1–10, 1966.

[98] G. Shechter, F. Devernay, E. Coste-Maniere, A. Quyyumi, and E.R. McVeigh. Three-dimensional motion tracking of coronary arteries in biplane cineangiograms. *IEEE Transactions on Medical Imaging*, 22:493–503, 2003.

[99] R. Shrager, D.K. Jones, S. Pajevic, P. Munson, and P.J. Basser. When is a gaussian displacement distribution adequate to describe water diffusion in tissues? In *Proceedings of the ISMRM Workshop on Diffusion MRI*, 2002.

[100] J. Sijbers, A.J. den Dekker, J. Van Audekerke, M. Verhoye, and D. Van Dyck. Estimation of the noise in magnitude mr images. *Magnetic Resonance Imaging*, 16:87–90, 1998.

[101] B.W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, New York, NY, 1986.

[102] W.R. Smyth. *Static and Dynamic Electricity*. MacGraw-Hill, New York, 1969.

[103] C. Studholme, R.T. Constable, and J.S. Duncan. Accurate alignment of functional epi data to anatomical mri using a physics-based distortion model. *IEEE Transactions on Medical Imaging*, 19:1115–1127, 2000.

[104] 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.

[105] C. Studholme, D.L.J. Hill, and D.J. Hawkes. Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Medical Physics*, 24:25–35, 1997.

[106] N.A. Thacker, A. Jackson, D. Moriarty, and E. Vokurka. Improved quality of re-sliced mr images using re-normalized sinc interpolation. *Journal of Magnetic Resonance Imaging*, 10:582–588, 1999.

[107] P. Thévenaz, T. Blu, and M. Unser. Image interpolation and resampling. In I.N. Bankman, editor, *Handbook of Medical Imaging, Processing and Analysis*, chapter 25, pages 393–420. Academic Press, San Diego CA, USA, 2000.

[108] P. Thévenaz, T. Blu, and M. Unser. Interpolation revisited. *IEEE Transactions on Medical Imaging*, 19(7):739–758, July 2000.

[109] J.-P Thirion. Image matching as a diffusion process: an analogy with maxwell's demons. *Medical Image Analysis*, 2:79–98, 1998.

[110] P. Thompson and A. Toga. A surface-based technique for warping 3-dimensional images of the brain. *IEEE Transactions on Medical Imaging*, 15:402–417, 1996.

[111] A. N. Tikhonov and V. A. Arsenin. *Solutions of Ill-Posed Problems*. Winston, Washington, DC, 1978.

[112] J. Tsao. Interpolation artifacts in multimodality image registration based on maximization of mutual information. *IEEE Transactions on Medical Imaging*, 22:854–964, 2003.

[113] D.S. Tuch, R.M. Weisskoff, J.W. Belliveau, and V.J. Wedeen. High angular resolution diffusion imaging of the human brain. In *Proceedings of the 7th Annual Meeting of ISMRM*, page 321, 1999.

[114] M. Unser, A. Aldroubi, and M. Eden. B-spline signal processing: Part I-theory. *IEEE Transactions on Signal Processing*, 41:821–833, 1993.

[115] M. Unser, A. Aldroubi, and M. Eden. B-spline signal processing: Part II-efficient design and applications. *IEEE Transactions on Signal Processing*, 41:834–848, 1993.

[116] M. Unser, A. Aldroubi, and M. Eden. Enlargement or reduction of digital images with minimum loss of information. *IEEE Transactions on Image Processing*, 4:247–257, 1995.

[117] M. Unser and P. Thévenaz. Stochastic sampling for computing the mutual information of two images. In *Proceedings of the Fifth International Workshop on Sampling Theory and Applications (SampTA'03)*, pages 102–109, Strobl, Austria, May 26-30, 2003.

[118] P. Viola and W.M. Wells III. Alignment by maximization of mutual information. In *Proceedings of the 5th International Conference on Computer Vision*, pages 16–23, 1995.

[119] V.J. Wedeen, T.G. Reese, D.S. Tuch, M.R. Weigel, J-G. Dou, R.M Weiskoff, and D. Chesler. Mapping fiber orientation spectra in cerebral white matter with Fourier trannsform diffusion MRI. In *Proceedings of the 8th Annual Meeting of ISMRM*, page 82, 2000.

[120] J. Weickert and G. Kühne. Fast methods for implicit active contour models. In S. Osher and N. Paragios, editors, *Geometric Level Set Methods in Imaging, Vision, and Graphics*. Springer, 2003.

[121] [9] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, Jr., R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. Hill, D. J. Hawkes, C. Studholme, J. B. Maintz, M. A. Viergever, G. Malandain, R. P. Woods, and et al. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of Computer Assisted Tomography*, 21:554–566, 1997.

[122] R.P. Woods, S.R. Cherry, and J.C. Mazziotta. Rapid automated algorithm for aligning and reslicing pet images. *Journal of Computer Assisted Tomography*, 16:620–633, 1992.

[123] R.P Woods, S. T. Grafton, T.J. Holmes, S.R. Cherry, and J.C. Mazziotta. Automated image registration: I. general methods and intrasubject, in-

tramodality validation. *Journal of Computer Assisted Tomography*, 22:139–152, 1998.

[124] R.P. Woods, S.T. Grafton, J.D.G. Watson, N.L. Sicotte, and J.C. Mazziotta. Automated image registration: II. intersubject validation on linear and nonlinear models. *Journal of Computer Assisted Tomography*, 22:153–165, 1998.

[125] C. Xu, D.L. Pham, and J.L. Prince. Medical image segmentation using deformable models. In J.M. Fitzpatric and M. Sonka, editors, *Handbook of Medical Imaging*. SPIE Press, 2000.

[126] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21:977–1000, 2003.