NORTHEASTERN UNIVERSITY

Graduate School of Engineering

Project Title: Investigation of a Volumetric Technique for Registration of Magnetic Resonance Images of the Human Brain.

Author: Patrick K. Rennich.

Department: Electrical and Computer Engineering.

Approved for Project Requirements of the Master of Science Degree:

Professor Eric L. Miller, Advisor	Date
Professor Dana H. Brooks	Date
	Date
	Date

Graduate School Notified of Acceptance:

Director, Graduate School	Date

Investigation of a Volumetric Technique for Registration of Magnetic Resonance Images of the Human Brain

A Project Presented

by

Patrick K. Rennich

to

The Electrical and Computer Engineering Department

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in the field of

Electrical Engineering

Northeastern University

Boston, Massachusetts

March 2001

Abstract

In order to more completely understand the living human brain additional computational tools are necessary to manipulate and analyze volumetric brain images. Comparisons between individual brains and across populations of brains are becoming increasingly important in characterizing both the anatomy and functioning of the brain. This work investigates a technique to register volumetric magnetic resonance images of the human brain in a nonlinear fashion allowing more meaningful comparisons between multiple individuals or between an individual brain and a standard brain atlas. The mathematics of 3-dimensional transformations are fully developed which provides a framework for construction of an objective function. This function is carefully designed to encourage a match while enforcing certain biological constraints, and minimization of the objective function in conjunction with a multi-resolution scheme brings about the registration between image volumes. Performance is analyzed which reveals shortcomings of the technique as well as providing insight into the operation of the algorithm. Potential modifications to rectify the problems are also discussed.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

During the past several decades there has been considerable interest in understanding the structure and functioning of the human brain. In the 1970s and 1980s the advent and development of minimally-invasive intracranial imaging techniques such as X-ray computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) has provided clinicians an invaluable view of the living brain. These imaging techniques have also enabled detailed study and analysis of cerebral anatomy and functioning by neuroscientists. In fact, the progress and promise of brain research prompted President George Bush in 1990 to proclaim the coming decade as "the decade of the brain" [1].

The speed of research has been considerable, but without direct access to the living brain through invasive means, the body of knowledge of human brain functioning has been slow to expand. Considering what is known about non-human primate brains such as monkeys in which invasive techniques are routinely utilized, comparatively little is known about the human brain [2, 3, 4, 5, 6, 7]. Perhaps the most telling commentary on the current state of knowledge of human cortex is that many contemporary imaging studies still report their

results using cortical area numbers from Brodmann's landmark 1909 parcellation of the human cortex. Full access to the monkey brain has propelled research in non-human primates far beyond what was identified by Brodmann, yet Brodmann based areas are still used for human studies.

Clearly, more sophisticated and ethically acceptable (i.e. non-invasive) tools are needed for expanding the body of knowledge regarding the living human brain. Intracranial imaging techniques can provide this needed insight, but more work is necessary in developing and refining the computational and analytical tools associated with brain imaging. In particular, an accurate characterization of brain structure and functioning across a group of individuals is still difficult to obtain. There are significant variations in the anatomical structure of brains between individuals which must be accounted for in some fashion before any meaningful comparisons of function can be made. A robust registration algorithm is a critical component of any research which seeks to objectively compare brain structure or functioning across multiple individuals.

The most common technique for accommodating differences between individuals is to register each brain dataset in the Talairach stereotaxic coordinate system [8] and then perform the comparison. This Cartesian coordinate system is based upon specific subcortical landmarks within the brain, and registration of the image volumes is then accomplished by finding the appropriate rotation, translation, and scaling of the axes in a piecewise linear fashion [8]. These transformations, however, fail to account for even large variations in brain anatomy particularly for cortical structures. Techniques which utilize additional nonlinear transformations and local scalings are necessary to accommodate these variations in normal (and diseased) brains.

These techniques which go beyond simple linear transformations are known as warping algorithms and their development is an active area of research. A wide variety of approaches currently exist which indicates the importance and need for these algorithms [9]. Most incorporate global scalings, linear, nonlinear, and local deformations, and utilize both

specific anatomic features as well as raw imaging data. However, a robust, automatic algorithm which achieves a precise warp (or match) between individuals or between an individual and a standard atlas has yet to emerge.

This work investigates a new approach to volumetric warping of the imaged human brain with the goals of being automatic, robust, and precise. The MRI modality which is arguably one of the most powerful for general investigation of the living human brain [10] is utilized for this work as it provides high resolution images of both structure and functioning in a completely non-invasive manner.

1.2 Applications of Brain Warping

Establishing a biologically viable voxel-to-voxel correspondence between an individual brain volume and either another individual or a composite brain atlas has a wide range of applications both clinically and in the research arena. Anatomic structures which correspond between individuals are called homologous, and they often exhibit a natural variability. For example, the ears of two individuals are homologous, yet are certainly distinct in their particular structure. Characterizing the variability of brain structure is useful in a number of contexts either to fully understand normal biologic variation or to minimize its effects as described in the applications below.

If the correspondence between a collection of brains were known, this information could be used to easily construct a statistical brain atlas which characterizes the global and local structural variations present within this collection. The ability to characterize particular classes of brains in this way has significant diagnostic power. For example, it would then be possible to compare a brain of an unknown class with the class of normal brains and quantitatively determine their similarity in a statistical sense. Comparisons of this type could quickly isolate and identify areas of the unknown brain which exceeded normal variation such as tumors and other structural pathologies. Used in a clinical setting,

this may not replace the physician's diagnosis, but would be a useful diagnostic aid for her.

From a research perspective, the ability to determine the warping between an individual brain volume and an established standard volume would allow researchers to precisely specify regions of the brain. The need for a standard way to do this has been a driving force behind the use of the popular Talairach stereotaxic coordinate system. One drawback of the Talairach system, however, is its inability to account for local structural variations– especially toward the outer regions of the brain where the variations are typically the greatest. Nonlinear warping algorithms have been developed to address this issue by allowing for localized deformations. With a robust and reliable warping technique scientists in any area of brain research can precisely reference very specific regions of the brain in terms of an established standard brain. Uniformity and precision are fundamental components of a reference system used to communicate research procedures and report results, and the lack of a system with such components has been a long standing challenge for researchers.

Warping techniques can also be applied across time measuring the growth of a single brain. Observing brain development in a quantitative fashion by warping from one point in time to the next significantly enhances our understanding of the formation of the human brain [11]. In a similar way, warping can be used to study the deterioration of brain tissue over time associated with neurodegenerative diseases such as Parkinson's and Alzheimer's.

In addition to quantifying and accentuating the structural differences between brain volumes either across individuals or over time, it is also useful to use warping to minimize these differences. In experiments that aim to study brain functioning across a population, differences in anatomy between individuals are a confounding factor that should be accounted for. To perform these types of studies requires the substrate upon which the functioning occurs to be brought into register in some fashion. This substrate is the structural anatomy of the brain. A warp between two individuals which is obtained using anatomical structure can be applied to brain activation data enabling the functioning of these brains to be compared in a more meaningful sense.

It is important to note that no single technique can satisfy all the varied requirements demanded of brain warping. For example, model based algorithms which obtain very good matches of homologous structures but require manual labeling of specific features are not suitable in situations which require a large number of brains to be automatically warped. Furthermore, it is not particularly clear what constitutes a *good* warp. Arriving at a warp which *exactly* matches intensities is not difficult even for brains which exhibit features that are unreconcilable in a biological sense [12]. Clearly, though, this warp would be of limited usefulness. Establishing certain constraints under which the warp can occur becomes the true problem. These constraints are not easily determined, and even once they are stated in biological terms, translation of the constraints into a mathematical algorithm compounds the difficulty. The validity of the warp is a significant biological question which is not answerable in a general sense, and must be determined within the context of how and where the warp will be applied.

1.3 Project Overview and Organization

This project seeks to understand and evaluate an alternative approach to the problem of determining a correspondence between two brain volumes. Certain deficiencies of previous approaches are identified and addressed with this new technique, and an evaluation of its successes and failures are presented.

The following exposition presents the background, development, analysis, and conclusions associated with a newly developed 3-dimensional warping technique for MRI acquired brain structures. Chapter 2 provides a review of previous work in this field including the Talairach atlas as well as more sophisticated contemporary approaches. Chapter 3 discusses the technical component of the current algorithm with an emphasis on the associated mathematics. Finally, a performance analysis and conclusions are presented in Chapters 4 and 5 respectively.

Chapter 2

Literature Review

Warping algorithms can generally be divided into two categories: intensity based and model based. The former seek to automatically register brain volumes based directly upon the intensity of the voxels while remaining within certain biologically established constraints. The latter rely upon the accurate labeling of structural features to guide the warp. These categories are not mutually exclusive as some approaches make use of both the intensity and structural features.

This chapter discusses the relevant information that is available in the literature concerning these techniques. Due to its pervasive use in the neuroscience communities, it is also necessary to discuss the brain reference system first developed by Talairach et al. known as the Talairach stereotaxic atlas [8, 13, 14].

2.1 Talairach Stereotaxic Atlas

The Talairach atlas was originally developed in the late 1950s as a reference frame for stereotaxic neurosurgery [15]. Using a stereotaxic device which is rigidly attached to the head of the patient, neurosurgeons can precisely position their surgical equipment within the brain of the patient facilitating the efficient biopsy or removal of tumors, epileptic foci,

Figure 2.1: *Stereotaxic neurosurgical devices. (Stereoguide and Stereoadapter from Sandström Trade and Technology Inc.)*

and other pathologies. (Figure 2.1 shows an example of a stereotaxic device.) As a standard coordinate based reference frame the Talairach stereotaxic atlas provides a common and precise language which is superior to descriptive images for specifying specific regions within the brain [16]. This affords surgeons with knowledge of not only the space occupied by the pathologic process but also of the anatomic structures surrounding the area [8].

The availability and simplicity of this reference frame led to its use in other areas of brain research such as brain mapping where it has become the *de facto* standard for communicating research results. Additionally, the ability to place any brain within the Talairach coordinate system has led to its use as a registration method. With this technique two brains can be brought into register simply by determining the Talairach coordinates for each brain. The Talairach atlas is also applicable over a wide variety of imaging modalities enabling inter-modality registration. For these reasons, use of the Talairach brain atlas has become widespread.

The Talairach coordinate system is specified by first defining a line which passes through the superior edge of the anterior commissure (AC) and the inferior edge of the

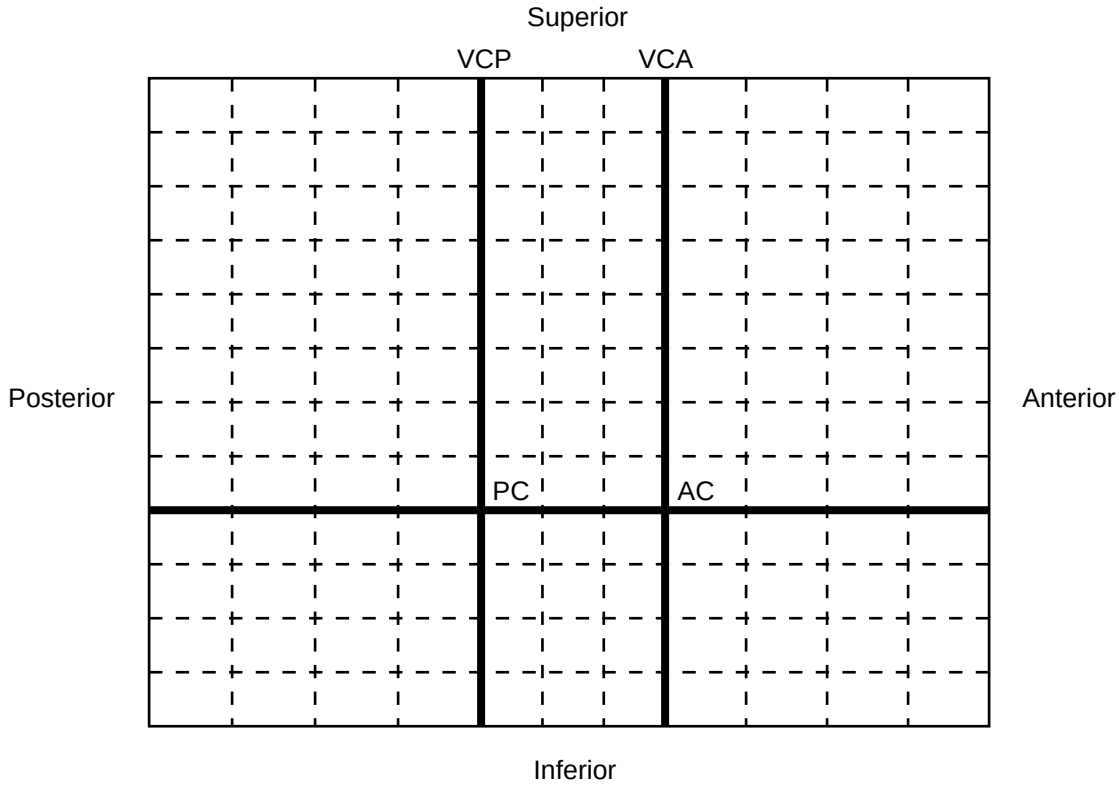

Figure 2.2: *Sagittal view of the Talairach proportional grid system.*

posterior commissure (PC) [8]. This line is known as the AC-PC line and is shown in Figure 2.2 as the bold horizontal line. The second step is to define the midplane as the interhemispheric sagittal plane. The horizontal plane is then defined as that which contains the AC-PC line and is perpendicular to the midplane. Third, a vertical line perpendicular to the horizontal plane and passing through the posterior margin of the anterior commissure specifies the vertical axis. A plane containing this line and perpendicular to the other two planes is known as the verticofrontal plane (VCA) [8]. The intersection point of these three planes specifies the origin of the coordinate system. A fourth plane is also defined which is parallel to the veritcofrontal plane but contains the posterior commissure (VCP).

With this system, each of the twelve sections are bounded by at least one anatomical landmark on the inside (AC, PC or both AC and PC). On the outside the sections are

bounded by the fullest extent of the brain in the principle directions defined by the coordinate system (anterior, posterior, superior, inferior, or lateral). Each particular brain extends by different amounts in these directions in absolute terms and so the Talairach coordinate system has the effect of linearly scaling each section based on the particular brain to which the coordinates are being applied.

In practice, registration of a brain volume with the Talairach coordinate system is automated by use of an intensity based affine transformation. The affine parameters are estimated by comparing the brain volume with a composite reference volume composed of brains previously registered to the Talairach coordinates. One such technique has been developed by Louis Collins et al. at the Montreal Neurological Institute [17].

This proportional coordinate system has proven to be quite useful for mapping the human brain, but there are some significant shortcomings. To bring two brains into register using the Talairach system essentially requires a global rigid rotation and translation followed by a piecewise linear scaling along each axis. This accounts for global and large structures of the brain as well as structures close to the anterior and posterior commissures. However, for structures distant from these commissures the mismatch can be significant (greater than several centimeters) [18, 19, 20, 21]. Many of the major sulci and gyri are not aligned with sufficient precision to perform meaningful structural or functional comparisons. This limitation has stimulated the development of a variety of warping techniques which are discussed in the next sections.

2.2 Intensity Based Algorithms

Intensity based techniques generally consist of two basic components: an index of similarity and a set of constraints which are used to help guide the generation of the warp. The similarity index is derived from the intensities of the two volumetric images, and is typically some form of a squared difference or correlation statistic. The constraints are

generally aimed at obtaining a one-to-one and continuous transformation. This type of transformation is known as a homeomorphism and is achieved either explicitly by direct evaluation of the warp field or implicitly through a smoothing of the warp field itself. This ensures that anatomic structures maintain their same relative positions before and after the transformation.

Most algorithms are iterative in nature letting the warp evolve in a series of smaller steps. In addition, some utilize a multi-resolution approach which seeks to simplify the problem by initializing the algorithm at coarse resolutions and then successively increasing the complexity of the warp. These techniques can potentially require a very large number of parameters (greater than 1×10^6) to effect the warp necessitating novel approaches for parameter estimation. Most also begin with a translation and linear transformation to initialize the large number of parameters. This initial operation is generally referred to in the literature as an affine transformation.

Some of the more prominent intensity based algorithms are those of Collins and Evans who use a smoothed piecewise linear approach known by the moniker ANIMAL [22] and Ashburner and Friston who make use of linear combinations of smooth basis warps [23]. Also of particular note is the surface based warping of Fischl and Dale [24, 25].

The ANIMAL algorithm of Collins and Evans has all the basic components discussed above where the similarity measure is a normalized correlation statistic of the intensities within a local neighborhood, the constraint is effected by smoothing the deformation vectors at each iteration, and a multi-resolution scheme is used. Starting at the coarsest level, the volumes are divided into cubical sections, and the optimal linear fit with respect to the correlation statistic is obtained for each section. The process is repeated at the next higher resolutions, ultimately building up a nonlinear deformation field in a piecewise linear fashion. Homeomorphic warps are not guaranteed, however, as the smoothing constraint does not explicitly restrict non-invertible transformations.

Ashburner and Friston take a different approach incorporating a set of smooth warps

composed of low frequency 3-dimensional discrete cosine transform basis functions. These 'basis warps' are then combined linearly to construct a full warp field which maximizes the *a posterior* probability of being a correct warp. This MAP scheme incorporates both the voxel intensities as well as prior probabilities of the warp parameters. These priors are not based upon empirical evaluation of a large set of subjects but upon the so-called membrane energy of the warp field [26]. The use of low frequency 'basis warps' precludes high resolution matching of the voxel intensities. However, the highest frequency can be adjusted to suit the specific application.

While not strictly an approach to volumetric warping, the surface matching algorithm of Fischl and Dale is noteworthy due to its similarity in structure to the algorithm currently being investigated. Once the surface of the cerebral hemispheres has been identified and mapped into a spherical coordinate system, the warp is constructed by minimization of an objective function. This objective function is composed of a curvature matching term to effect the registration, an areal term to prevent folding, and a metric distortion term to encourage the preservation of surface distances. Weighted appropriately, minimization of these terms constructs a transformation of the cerebral cortex based upon the sulcal and gyral patterns.

2.3 Model Based Algorithms

Model based algorithms rely upon identification of specific features or landmarks associated with the brain structure. Of course, these features must be observable under the particular imaging modality in use and be present in all the brains under consideration. These homologous features generally consist of discrete points, curves and contours, and/or surfaces, and are usually identified by manual inspection of the image volume. Manual labeling can be subjective and is generally time consuming which in some situations decreases

the attractiveness of these approaches. Additionally, there is a paucity of readily identifiable discrete points in the brain from which to anchor the warp [12]. Curved features are also somewhat difficult to identify especially in an automatic fashion. Surfaces, on the other hand, have been shown to be more readily extracted [25].

Once these landmarks have been determined, the goal is to find the transformation which superposes the landmarks in the source image with the homologous landmarks in the reference image. For point landmarks the transformation is straight-forward where homologous points are known because they have been identified and the remaining volume is simply forced to be smooth through some interpolation scheme. More complex interpolation is used for curves and surfaces such as with the fluid model of Thompson and Toga [27].

Other algorithms such as that of Gérard Subsol make use of curves obtained by following sulcal and gyral contours [28]. After an initial intensity based affine transformation these contours are matched with a nearest neighbor algorithm. The transformation of the remaining volume is then determined by using a spline based technique.

Another approach used by Downs et al. first obtains the 3-dimensional convex hull determined by the brain surface [15]. This is matched to a template convex hull by proportionately scaling the volume radially from the coordinate system origin located centrally within the brain.

Chapter 3

Warping Architecture

3.1 Mathematical Preliminaries

This section lays out the mathematical underpinnings for the current warping algorithm. The notational conventions are described Subsection 3.1.1, and the fundamental transformation mathematics upon which these techniques are based are described in Subsection 3.1.2. Additionally, a transformation construct known as the Jacobian is fully developed mathematically including a geometrical description in Subsection 3.1.3.

3.1.1 Notation

The canonical dimensions in the 3-dimensional space R^3 are denoted by the numbers 1, 2, and 3 and when associated with a symbol are written as a subscript. Vectors are represented by lowercase bold face symbols. Discrete sets of vectors are represented by the uppercase version of the same bold face symbol, and individual components of the set are indicated by a parenthesized superscript. An example of this notation follows:

x_1 : a continuous variable in the 1-axis,

$\mathbf{x}$: a continous vector,

$\mathbf{D}$: a discrete set of vectors,

$\mathbf{d}^{(i)}$: the i^{th} vector of the discret set $\mathbf{D}$,

$d_2^{(i)}$: the 2-axis component of $\mathbf{d}^{(i)}$.

The remaining notational conventions should be evident from the surrounding discussion.

3.1.2 Warps and Transformations

A volumetric warp is described mathematically as a transformation (or mapping) between two coordinate spaces in R^3, and we denote such a mapping as $F : R^3 \rightarrow R^3$ where F is a vector-valued function of a vector. If we let the 3-D vector $\mathbf{x}$ represent coordinates in R^3 (x_1, x_2, x_3) we can formalize the transformation F as follows:

$$F(\mathbf{x}) = \Big(u_1(\mathbf{x}), u_2(\mathbf{x}), u_3(\mathbf{x})\Big)$$

where u_1, u_2, and u_3 are called the component functions of the transformation F. The resulting 3-D coordinate (u_1, u_2, u_3) will be referred to as the vector $\mathbf{u}$ for the remainder of this thesis.

Mappings that are used as warps between homologous biological structures (the brain in the current case) should be both continuous and homeomorphic. The mathematical expression of these two properties are developed in the following paragraphs.

For a real-valued function of a real variable $f : R \rightarrow R$, continuity is defined as follows. At a point x_0, if given an $\varepsilon > 0$ there exists a $\delta > 0$ such that if $|x - x_0| < \delta$ then $|f(x) - f(x_0)| < \varepsilon$. This concept is extended in a straightforward manner to 3-D mappings. Let $\mathbf{x}$ and $\mathbf{y}$ be coordinates in R^3. A map $F : R^3 \rightarrow R^3$ is continuous at a point $\mathbf{x}$ if given an $\varepsilon > 0$ there exists a $\delta > 0$ such that if $\|\mathbf{x} - \mathbf{y}\|_2 < \delta$ then $\|F(\mathbf{y}) - F(\mathbf{x})\|_2 < \varepsilon$ (the notation $\|\cdot\|_2$ represents the 2-norm) [29]. This concept is illustrated further in Figure 3.1.

For an arbitrary set A contained in R^3, the mapping $F : A \subset R^3 \rightarrow R^3$ is said to be a homeomorphism onto $F(A)$ if F is one-to-one and if the inverse $F^{-1} : F(A) \subset R^3 \rightarrow R^3$

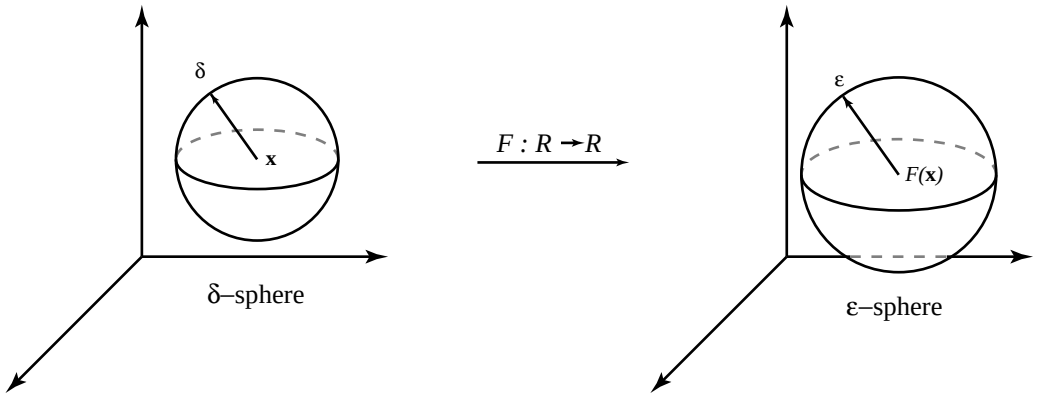

Figure 3.1: *Continuity of a 3-D mapping: If the point $F(\mathbf{y})$ exists inside the ε-sphere, then the point* $\mathbf{y}$ *must exist inside the δ-sphere for arbitrarily small ε.*

is continuous [29]. Intuitively in 2-D a homeomorphism applies a transformation without tearing or folding of the space in R^2. A key property of a mapping that is relied upon heavily in the warping algorithm is that the Jacobian matrix of a homeomorphic mapping is positive semi-definite everywhere that the mapping is defined (this concept is more fully described in Subsection 3.1.3).

As with the concept of continuity, differentiation of a transformation is easily obtained by expanding the usual notion of differentiation of a function of a single variable to 3-D mapping functions. The partial derivative of the mapping F with respect to, say, x_1 is a vector of the partial derivatives of each of the component functions of F:

$$\frac{\partial F}{\partial x_1} = \left(\frac{\partial u_1}{\partial x_1}, \frac{\partial u_2}{\partial x_1}, \frac{\partial u_3}{\partial x_1} \right) = \frac{\partial \mathbf{u}}{\partial x_1}.$$

Now, when the partial derivative of the mapping F with respect to the vector $\mathbf{x}$ is desired, we construct a matrix of the partial derivatives of the component functions as follows:

$$\frac{\partial F}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \frac{\partial u_1}{\partial x_3} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} & \frac{\partial u_2}{\partial x_3} \\ \frac{\partial u_3}{\partial x_1} & \frac{\partial u_3}{\partial x_2} & \frac{\partial u_3}{\partial x_3} \end{bmatrix} = \begin{bmatrix} | & | & | \\ \frac{\partial \mathbf{u}}{\partial x_1} & \frac{\partial \mathbf{u}}{\partial x_2} & \frac{\partial \mathbf{u}}{\partial x_3} \\ | & | & | \end{bmatrix}.$$

This matrix is known as the Jacobian of the transformation, and the determinant of this matrix will play an important role in the development of the warping algorithm. As shown

in Equation (3.1), the determinant is computed by using the triple-product where $\times$ denotes the vector outer-product, $\cdot$ represents the vector inner-product, and J denotes the Jacobian determinant.

$$\begin{aligned} J &= \left(\frac{\partial u_1}{\partial x_1}, \frac{\partial u_2}{\partial x_1}, \frac{\partial u_3}{\partial x_1}\right) \times \left(\frac{\partial u_1}{\partial x_2}, \frac{\partial u_2}{\partial x_2}, \frac{\partial u_3}{\partial x_2}\right) \cdot \left(\frac{\partial u_1}{\partial x_3}, \frac{\partial u_2}{\partial x_3}, \frac{\partial u_3}{\partial x_3}\right) \\ &= \frac{\partial \mathbf{u}}{\partial x_1} \times \frac{\partial \mathbf{u}}{\partial x_2} \cdot \frac{\partial \mathbf{u}}{\partial x_3} \end{aligned} \tag{3.1}$$

Further discussion of the geometrical interpretation of the Jacobian is contained in Subsection 3.1.3.

3.1.3 Jacobian Geometry

The Jacobian matrix associated with the transformation function plays an important role in the regularization of the warp field estimate and deserves further discussion. A geometric interpretation and details of the mathematical implementation are presented in the next few paragraphs.

The determinant of the Jacobian matrix (referred to as simply the Jacobian and denoted as J) computed at a specific coordinate indicates the amount of contraction or expansion that the transformation induces in the neighborhood of that coordinate. The properties of the Jacobian are summarized below:

$$J \begin{cases} > 1 & \Rightarrow \text{local expansion,} & \text{Property (1)} \\ = 1 & \Rightarrow \text{volumetric preservation,} & \text{Property (2)} \\ < 1 & \Rightarrow \text{local contraction,} & \text{Property (3)} \\ = 0 & \Rightarrow \text{contraction of the local volume to a point,} & \text{Property (4)} \\ < 0 & \Rightarrow \text{volumetric inversion.} & \text{Property (5)} \end{cases}$$

To understand why the Jacobian has these properties consider the differential volume elements as shown in Figure 3.2. Part (a) of this figure shows a cubical element which is transformed by a mapping $F : R^3 \rightarrow R^3$ into the element shown in Part (b) of the same

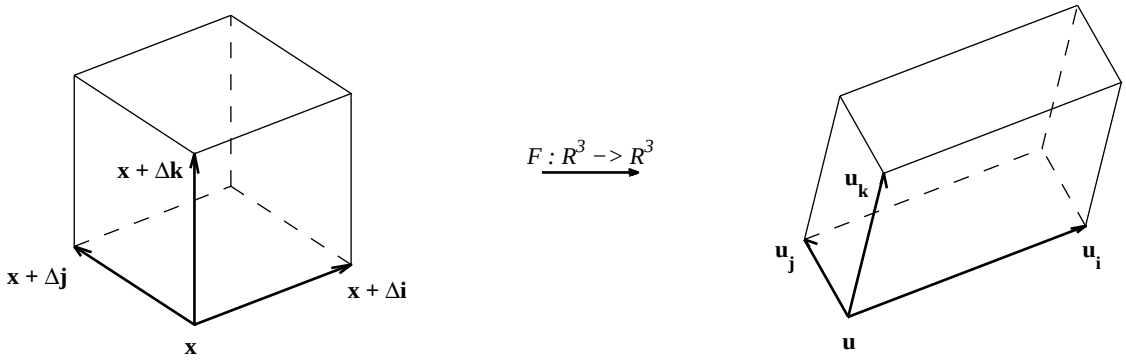

Figure 3.2: *Differential volume element: (a) original, and (b) after transformation under the hypothetical mapping* $F : R^3 \rightarrow R^3$.

figure. Given the dimensions of this differential element in Figure 3.2(a) $(\Delta\hat{\mathbf{i}}, \Delta\hat{\mathbf{j}}, \Delta\hat{\mathbf{k}})$ we can express the volume as $V = \Delta^3$. The transformation of this element under the mapping F is found by computing the image of the four coordinates $\mathbf{x}$, $\mathbf{x}+\Delta\hat{\mathbf{i}}$, $\mathbf{x}+\Delta\hat{\mathbf{j}}$, $\mathbf{x}+\Delta\hat{\mathbf{k}}$ as shown below:

$$\begin{aligned}
F(\mathbf{x}) &= \Big(u_1(\mathbf{x}), u_2(\mathbf{x}), u_3(\mathbf{x})\Big) &&= \mathbf{u}, \\
F(\mathbf{x}+\Delta\hat{\mathbf{i}}) &= \Big(u_1(\mathbf{x}+\Delta\hat{\mathbf{i}}), u_2(\mathbf{x}+\Delta\hat{\mathbf{i}}), u_3(\mathbf{x}+\Delta\hat{\mathbf{i}})\Big) &&= \mathbf{u}_i, \\
F(\mathbf{x}+\Delta\hat{\mathbf{j}}) &= \Big(u_1(\mathbf{x}+\Delta\hat{\mathbf{j}}), u_2(\mathbf{x}+\Delta\hat{\mathbf{j}}), u_3(\mathbf{x}+\Delta\hat{\mathbf{j}})\Big) &&= \mathbf{u}_j, \\
F(\mathbf{x}+\Delta\hat{\mathbf{k}}) &= \Big(u_1(\mathbf{x}+\Delta\hat{\mathbf{k}}), u_2(\mathbf{x}+\Delta\hat{\mathbf{k}}), u_3(\mathbf{x}+\Delta\hat{\mathbf{k}})\Big) &&= \mathbf{u}_k.
\end{aligned}$$

After transformation we now have a parallelopiped shaped differential element as depicted in Figure 3.2(b). To compute the volume of this parallelopiped (V') we can use the triple-product formula as shown below [30]:

$$V' = (\mathbf{u}_i - \mathbf{u}) \times (\mathbf{u}_j - \mathbf{u}) \cdot (\mathbf{u}_k - \mathbf{u}).$$

If we now take the ratio of the transformed volume V' and the original volume V as

is shown in Equation (3.2), we will have a quantity with the exact same properties as the Jacobian.

$$\frac{V'}{V} = \frac{(\mathbf{u}_i - \mathbf{u}) \times (\mathbf{u}_j - \mathbf{u}) \cdot (\mathbf{u}_k - \mathbf{u})}{\Delta^3}. \tag{3.2}$$

Specifically, these properties are (corresponding to those on page 16): (1) if the transformation induces an expansion of the differential element then the ratio is *larger than unity*, (2) if the volume remains the same after transformation then the ratio is *equal to unity*, (3) if the transformation causes a contraction of the differential element then the ratio is *less than unity*, (4) if the transformation produces a singular point then the ratio is *equal to zero*, and (5) if the transformed differential element has a negative volume (volumetric inversion) then the ratio is *less than zero*.

To further explore the relationship between this ratio and the Jacobian we can rewrite Equation (3.2) as follows:

$$\frac{V'}{V} = \frac{(\mathbf{u}_i - \mathbf{u})}{\Delta} \times \frac{(\mathbf{u}_j - \mathbf{u})}{\Delta} \cdot \frac{(\mathbf{u}_k - \mathbf{u})}{\Delta}. \tag{3.3}$$

For the moment, looking only at the first term of the vector outer-product in Equation (3.3) and recalling the first two equations of Equation (3.1.3) we can expand this term as such

$$\frac{(\mathbf{u}_i - \mathbf{u})}{\Delta} = \left(\frac{u_1(\mathbf{x} + \Delta\hat{\mathbf{i}}) - u_1(\mathbf{x})}{\Delta}, \frac{u_2(\mathbf{x} + \Delta\hat{\mathbf{i}}) - u_2(\mathbf{x})}{\Delta}, \frac{u_3(\mathbf{x} + \Delta\hat{\mathbf{i}}) - u_3(\mathbf{x})}{\Delta} \right) \tag{3.4}$$

which we recognize as the discrete differential of $\mathbf{u}$ at the coordinate $\mathbf{x}$ along the 1-axis ($\hat{\mathbf{i}}$ direction). Taking the limit of this quantity as Δ tends to 0 we obtain the partial differential as shown:

$$\lim_{\Delta \to 0} \frac{(\mathbf{u}_i - \mathbf{u})}{\Delta} = \left(\frac{\partial u_1}{\partial x_1}, \frac{\partial u_2}{\partial x_1}, \frac{\partial u_3}{\partial x_1} \right) = \frac{\partial \mathbf{u}}{\partial x_1}. \tag{3.5}$$

Now returning to the other two terms on the right hand side of Equation (3.3) and proceeding in a similar fashion we have:

$$\begin{aligned} \lim_{\Delta \to 0} \frac{(\mathbf{u}_j - \mathbf{u})}{\Delta} &= \left(\frac{\partial u_1}{\partial x_2}, \frac{\partial u_2}{\partial x_2}, \frac{\partial u_3}{\partial x_2} \right) = \frac{\partial \mathbf{u}}{\partial x_2}, \\ \lim_{\Delta \to 0} \frac{(\mathbf{u}_k - \mathbf{u})}{\Delta} &= \left(\frac{\partial u_1}{\partial x_3}, \frac{\partial u_2}{\partial x_3}, \frac{\partial u_3}{\partial x_3} \right) = \frac{\partial \mathbf{u}}{\partial x_3}. \end{aligned} \tag{3.6}$$

Equations (3.5) and (3.6) allow us to find the limit of Equation (3.3) as Δ tends to 0 as shown here

$$\lim_{\Delta \to 0} \frac{V'}{V} = \frac{\partial \mathbf{u}}{\partial x_1} \times \frac{\partial \mathbf{u}}{\partial x_2} \cdot \frac{\partial \mathbf{u}}{\partial x_3}$$

which is identical to the Jacobian as computed in the second equation of Equations (3.1). This shows that in the limit as Δ tends to 0 the ratio of the differential volume before and after transformation is equivalent to the Jacobian matrix determinant.

This geometric interpretation provides insight into the desirable properties of this quantity. Of particular interest is Property (5) shown on page 16. A Jacobian that is less than zero means that at this particular point the volume has been inverted by the transformation such that the vectors $\frac{\partial \mathbf{u}}{\partial x_1}$, $\frac{\partial \mathbf{u}}{\partial x_2}$, and $\frac{\partial \mathbf{u}}{\partial x_3}$ are no longer arranged in right-hand order, and the inverse mapping is not continuous. This provides the motivation for the homeomorphic term of the cost function used in the warp optimization as will be shown in Subsection 3.4.3. By keeping the Jacobian determinant positive we maintain a homeomorphic mapping.

The Jacobian provides a good measure of the local volumetric distortion induced by the transformation. However, transformations can be constructed that maintain the local volume such that the Jacobian is unity but also induce severe changes in shape. Imagine a differential parallelopiped whose volume is equal to Δ^3 but whose shape is radically different that that of Figure 3.2(a). In this case, the Jacobian is greater than zero (in fact it is unity) and so has not violated any homeomorphic constraints, but the angular change exhibited can be just as non-biological as non-homeomorphisms. This is the primary motivation for an additional term of the objective function as will be discussed in Subsection 3.4.3.

3.2 Algorithm Overview

The overriding goals which guided the development of the current warping algorithm is that it provide an automatic, precise, robust, and objective method for establishing a viable correspondence between volumetric brain image sets. Algorithms which require manual

intervention preclude efficient processing of large numbers of image sets and also introduce issues of repeatability and objectivity. Different experts or even the same expert at different times can label an individual volume inconsistently leading to non-repeatable results. A properly designed automatic algorithm, on the other hand, is applied consistently and objectively across the datasets presented to it. The following paragraph and figure present an overview of such an algorithm.

As shown in the diagrammatic overview of Figure 3.3, the warping algorithm consists of multiple steps. First, the algorithm is initialized with an affine registration between the source (S) and reference (R) image volumes. The result of this step is a vector field which represents a discrete sampling of the 3-D mapping function that achieves the optimal affine registration (Section 3.3). This vector field is then used to initialize the nonlinear portion of the algorithm which begins by subsampling S and R into what is known as a Gaussian pyramid (Section 3.4.1). The process then proceeds by iteratively applying the registration and evaluating the objective function which is minimized by utilizing the gradient information in a steepest descent fashion (Sections 3.4.3 and 3.4.4). This step constitutes the inner nonlinear optimization loop. Once the objective function asymptotes, the outer loop is executed which increases the resolution of the source and reference volumes by stepping up one level of the Gaussian pyramid. The registration vector field is also interpolated to match the resolution of the image volumes at this point, and the inner optimization loop is then again executed. The algorithm continues in this fashion until the final resolution is reached.

This algorithm is similar to the surface based work of Fischl and Dale [24, 25] in that an objective function with three terms is utilized to match intensity, prevent folding, and preserve volume. No geometric features are used explicitly, and hence, this algorithm falls within the intensity based category.

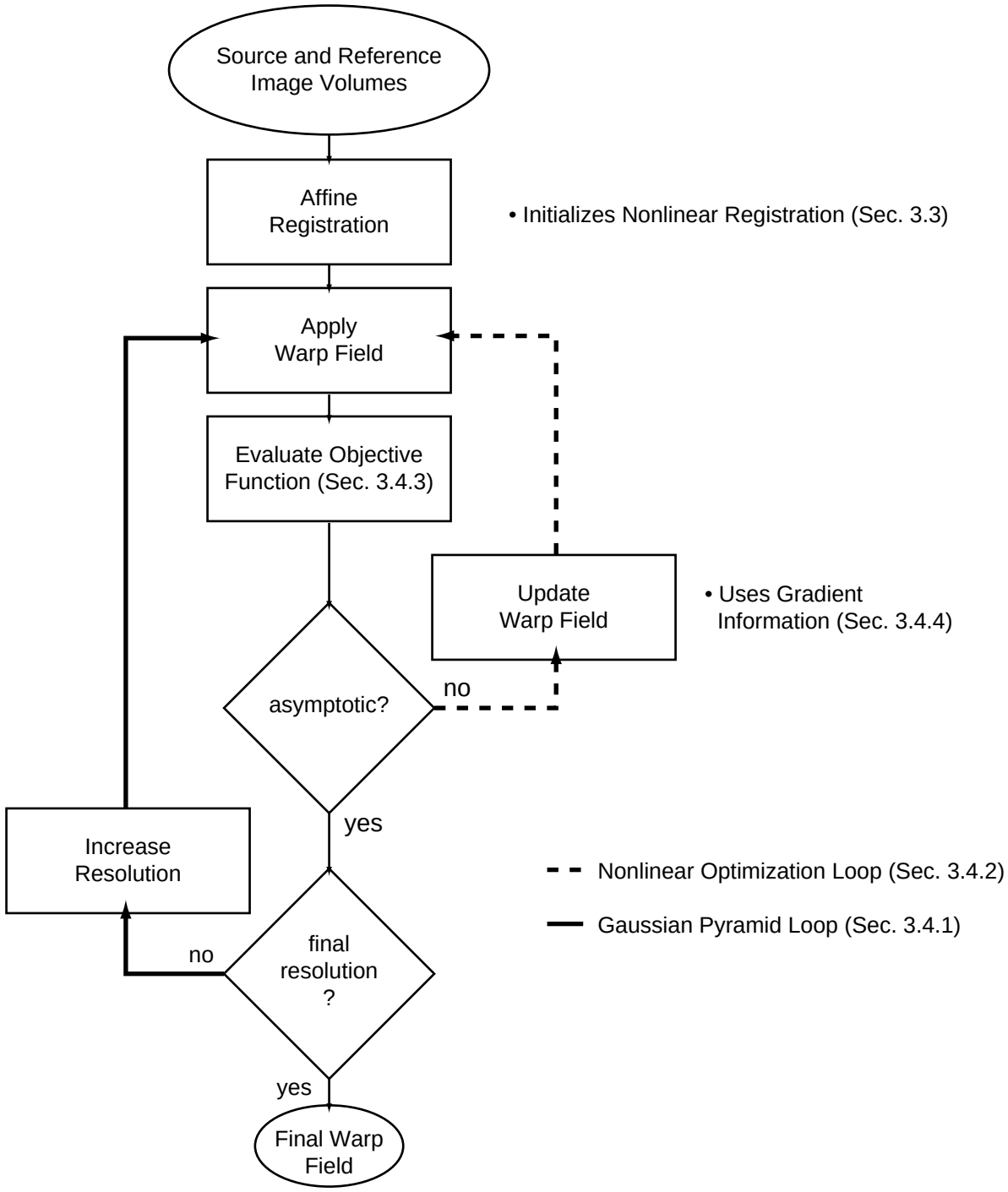

Figure 3.3: *Diagrammatic overview of the algorithm flow.*

3.3 Affine Registration

As is standard practice with most of the intensity based warping techniques reviewed in the previous chapter, the algorithm is initialized with an optimal affine transformation. The objective of this initial step is to bring the source and reference image volumes (S and R respectively) into an approximate global alignment before applying the nonlinear portion of the algorithm which operates both globally and locally. The affine transformation consists of rotation-, scale-, shear-, and translation-type transformations applied globally to the source volume resulting in a coarse alignment between the source and reference images. This provides a suitable starting point for the nonlinear alignment scheme that follows subsequently.

To further elaborate on the 3-D affine transformation, it is parameterized by a 4×4 homogeneous transformation matrix M which maps 3-D homogeneous coordinates $\mathbf{x}$ in the source image to coordinates $\mathbf{u}$ in the reference image as shown in the following equation [31]:

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ 1 \end{bmatrix} = \begin{bmatrix} m_1 & m_4 & m_7 & m_{10} \\ m_2 & m_5 & m_8 & m_{11} \\ m_3 & m_6 & m_9 & m_{12} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix}$$

More compactly, this is written as $\mathbf{u} = M\mathbf{x}$, and we see this to be consistent with the mathematical formalism presented in Subsection 3.1.2 by expressing the component functions of the transformation explicitly:

$$\begin{aligned} u_1(x_1,x_2,x_3) &= m_1x_1 + m_4x_2 + m_7x_3 + m_{10}, \\ u_2(x_1,x_2,x_3) &= m_2x_1 + m_5x_2 + m_8x_3 + m_{11}, \\ u_3(x_1,x_2,x_3) &= m_3x_1 + m_6x_2 + m_9x_3 + m_{12}. \end{aligned}$$

where u_1, u_2, and u_3 are the component functions of this transformation. (The last equation of this system is simply the tautology: $1 = 0x_1 + 0x_2 + 0x_3 + 1$.) Taken together, the

parameters m_1 through m_9 of the transformation matrix M, are responsible for the rotation, scale, and shear components of the affine transformation. The remaining parameters m_{10}, m_{11}, and m_{12} apply translation along the 1-, 2-, and 3-axes respectively.

These twelve affine parameters (m_1 to m_{12}) are estimated by minimizing the sum of squared errors between the intensities of the reference and the transformed source image volumes and across all voxels (V). This forms the objective function $C(M)$ [Equation (3.7)] which is used in conjunction with a quasi-newton optimization routine.

$$C(M) = \frac{1}{V}\sum_{i=1}^{V}\left[S(\mathbf{u}^{(i)}) - R(\mathbf{x}^{(i)})\right]^2 \tag{3.7}$$

Of course, the gradient of the objective function is required when using this type of optimization scheme. The gradient taken with respect to the parameters m_1 through m_{12} is derived as follows:

$$\nabla C(M) = \frac{1}{V}\sum_{i=1}^{V} 2\left[S(\mathbf{u}^{(i)}) - R(\mathbf{x}^{(i)})\right]\nabla S(\mathbf{u}^{(i)})$$

where $\mathbf{u}^{(i)} = M\mathbf{x}^{(i)}$. The full gradient $\nabla S(\mathbf{u}^{(i)})$ is constructed by taking the partial derivative of $S(\mathbf{u}^{(i)})$ with respect to each parameter using the chain rule as follows:

$$\frac{\partial S(\mathbf{u}^{(i)})}{\partial m_j} = \frac{\partial S(\mathbf{u}^{(i)})}{\partial u_1}\frac{\partial u_1}{\partial m_j} + \frac{\partial S(\mathbf{u}^{(i)})}{\partial u_2}\frac{\partial u_2}{\partial m_j} + \frac{\partial S(\mathbf{u}^{(i)})}{\partial u_3}\frac{\partial u_3}{\partial m_j}, \tag{3.8}$$

where m_j represents the j^{th} affine parameter. The partial derivative $\left(\frac{\partial S(\mathbf{u}^{(i)})}{\partial u_1}, \frac{\partial S(\mathbf{u}^{(i)})}{\partial u_2}, \frac{\partial S(\mathbf{u}^{(i)})}{\partial u_3}\right)$ is simply the intensity gradient of the volume S at the point $\mathbf{u}^{(i)}$ computed using a finite difference technique. When implemented, Equation (3.8) is actually simplified somewhat due to the fact that for any single m_j only one of the terms is ever non-zero [12].

This gradient is then employed as part of the quasi-newton optimization to obtain an estimate of M which fully specifies the affine transformation.

3.4 Nonlinear Warp Optimization

The warp field to be determined is a discrete sampling of the transformation function F that maps coordinates in R^3 (the $\mathbf{x}^{(i)}$s) to transformed coordinates which are again in R^3 (the $\mathbf{u}^{(i)}$s). Formally, this type of transformation is expressed as $F : R^3 \rightarrow R^3$ as discussed in Subsection 3.1.2. For each sampling point in the 3-D sampling grid the component functions of this transformation are written as follows:

$$\begin{aligned} u_1(x_1,x_2,x_3) &= x_1 + d_1, \\ u_2(x_1,x_2,x_3) &= x_2 + d_2, \\ u_3(x_1,x_2,x_3) &= x_3 + d_3, \end{aligned}$$

where, for simplicity, we have dropped the superscript (i) on each symbol. Of course, d_1, d_2, and d_3 for each sampling point are the quantities to be estimated and represent the displacement vector $\mathbf{d}$ for a particular location. More precisely we denote the displacement vector at the i^{th} location as $\mathbf{d}^{(i)}$ and the full set of $\mathbf{d}^{(i)}$s is denoted as $\mathbf{D}$. Exactly how the estimation of $\mathbf{D}$ is performed is described next.

3.4.1 Multiresolution Components

After the affine component of the transformation has been estimated, a multiresolution approach is used to obtain the global and local nonlinear deformations necessary to achieve a match between the reference and source MR volumes. This multiresolution scheme known as a Gaussian pyramid utilizes subsampled versions of the volumetric image set. A notional depiction of this scheme is shown in Figure 3.4. The pyramid is built by using the full-resolution ($256\times256\times256$) volume as the foundation, and each subsequent level is constructed by appropriately subsampling the previous level. In the current algorithm, a Gaussian generating kernel as specified in [32] is used to low-pass filter the volume before directly subsampling by a factor of two in each dimension using trilinear interpolation.

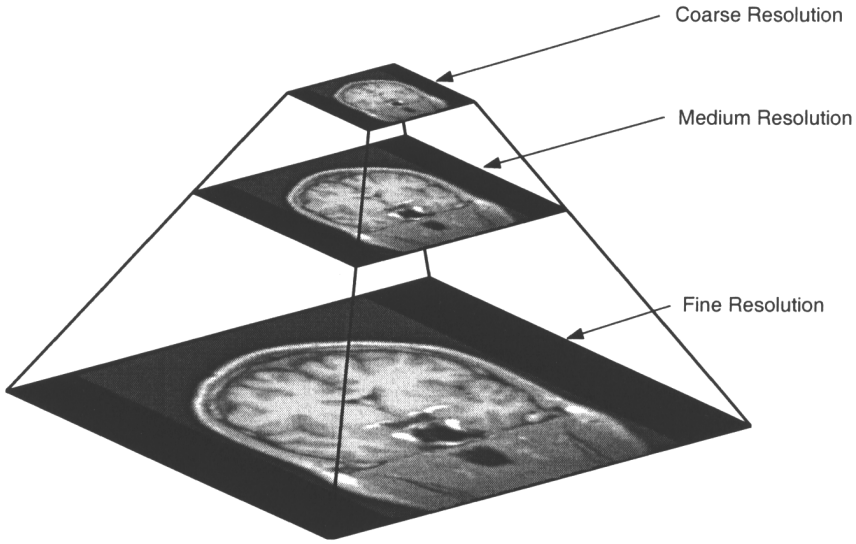

Figure 3.4: *Notional depiction of a Gaussian pyramid.*

The pyramid is completed once the lowest resolution level (currently set to $32\times32\times32$) is constructed which represents the top of the pyramid. Example slices from a Gaussian pyramid constructed in this way are shown in Figure 3.5. The nonlinear optimization procedure begins at the top of the pyramid and then progresses downward through the levels of increasing resolution as is outlined in the next subsection.

3.4.2 Optimization Structure

An iterative optimization procedure is utilized which is based upon an objective function constructed to achieve two major goals: first, the objective function should encourage a match between the MR intensities of the two volumes being considered, and second, it should prevent any biologically unreasonable transformations from occurring.

Beginning with the lowest resolution volumes at the top of the Gaussian pyramid, the optimization routine iteratively refines the warp field estimate by computing the gradient of the objective function and applying the result to the current estimate in a steepest descent

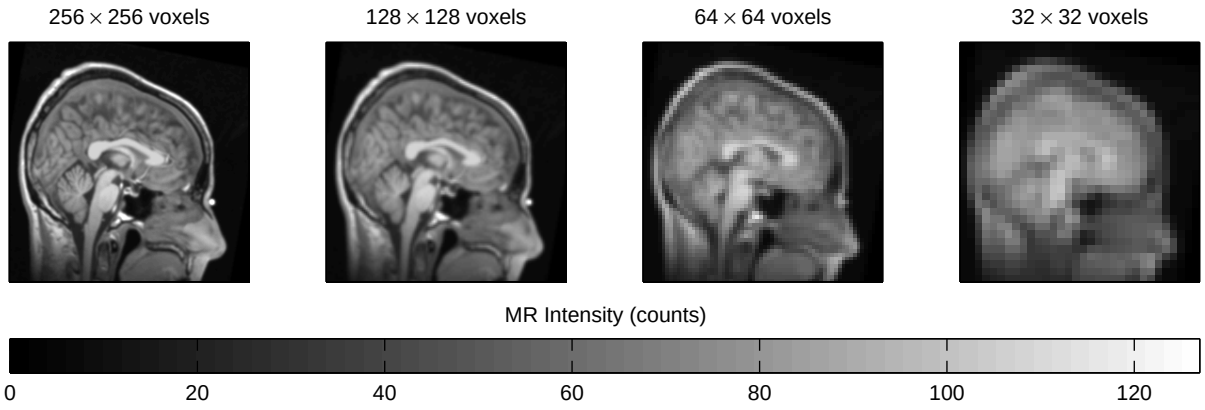

Figure 3.5: *A mid-sagittal slice of a brain volume represented at each level of a Gaussian pyramid.*

fashion. The objective function itself is also evaluated at each iteration to determine the progress of the procedure. If certain termination criteria are met, the optimization at the current level of the pyramid stops. The most recent estimate of the warp field is then interpolated by a factor of two so that the optimization can then begin at the next lower level (but higher resolution) of the Gaussian pyramid. Again, a trilinear interpolation technique is used to upsample the warp field.

The algorithm proceeds in this way until the base level (highest resolution) of the Gaussian pyramid is reached. At this resolution ($256 \times 256 \times 256$) the warp field is not interpolated due to computer memory constraints. Hence, each vector of the warp field now represents the transformation of a block of eight voxels. Other than this difference the optimization continues as before until the termination criteria are again satisfied. The resulting warp field represents a discrete sampling of the transformation function that registers the source volume with the reference volume.

3.4.3 Objective Function

At the heart of the optimization procedure lies the objective function. As noted before, the objective function serves a dual purpose: to encourage a match between intensities and to regularize or constrain the warp such that it makes sense biologically. The first goal is achieved through a similarity measure, and the second goal is accomplished by the combination of a metric distortion term and a term designed to severely penalize inversions of the volume. The general form of the full objective function $C(\mathbf{D})$ is expressed in Equation (3.9).

$$C(\mathbf{D}) = \lambda_1 C_1(\mathbf{D}) + \lambda_2 C_2(\mathbf{D}) + \lambda_3 C_3(\mathbf{D}) \tag{3.9}$$

The first term $C_1(\mathbf{D})$ is known as the similarity measure and computes the mean squared error between the voxel intensities of the reference volume (R) and the warped source volume (S). The next two terms compose the regularization component where $C_2(\mathbf{D})$ is the homeomorphic preservation term, and $C_3(\mathbf{D})$ is the metric distortion term. The first term of the regularization component $C_2(\mathbf{D})$ maintains a homeomorphism by computing the Jacobian determinant at all points in the warp field and applying a large cost to those that tend negatively. The next term $C_3(\mathbf{D})$ controls metric distortions by minimizing the mean squared change in the relative displacements among sampling points within some neighborhood. Each term is associated with a corresponding coefficient λ_{1-3} which determines its strength relative to the other terms. The set of vectors $\mathbf{D}$ represents the warp field which is to be estimated.

The similarity measure is more formally stated as

$$C_1(\mathbf{D}) = \frac{1}{V} \sum_{i=1}^{V} \frac{\left(R(\mathbf{x}^{(i)}) - S(\mathbf{x}^{(i)} + \mathbf{d}^{(i)}) \right)^2}{\Omega^2(\mathbf{x}^{(i)})}, \tag{3.10}$$

where R is the continuous 3-dimensional intensity function representing the reference volume, S is the continuous 3-dimensional intensity function representing the source volume,

and Ω^2 is an estimate of the intensity variance from brain to brain as a function of 3-dimensional space. (The functions R, S, and Ω^2 are more fully described in Section 4.1.) The vector $\mathbf{x}^{(i)}$ is the i^{th} sampling position on a regularly spaced 3-D sampling lattice, and $\mathbf{d}^{(i)}$ is the displacement vector associated with this sampling position. The set of all $\mathbf{d}^{(i)}$s is known collectively as the warp field and denoted simply as $\mathbf{D}$, and V is the total number of discrete points at which these volumes are sampled.

This measure computes the squared difference between voxel intensities in the reference and the corresponding voxel intensities of the warped source. This provides an index of similarity, but not all intensity differences should be treated equally. Some regions of the brain are expected to exhibit a higher level of variability from brain to brain, and this is accounted for through normalization by an estimate of the variance $\Omega^2(\mathbf{x}^{(i)})$ at each point in the brain volume. Typical examples of R and S are given in Section 4.1 as well as a more complete description of how Ω^2 is constructed and used.

Implicit in this formulation of the similarity term is the assumption of pre-calibrated intensity levels where the levels corresponding to certain tissue classes are consistent from reference (R) to source (S). This form, however, does not necessarily preclude uncalibrated intensities. If the calibration error between the source and reference is known to be at least linear then a simple scaling factor can be included with either the source or reference functions and estimated in conjunction with the warp parameters. For simplicity and since pre-calibrated image sets are readily available this modification was not included. Section 4.1 contains further details on the calibration technique applied to all brain volumes for this work.

Minimization of this term alone will produce a warp field which achieves a match between the source and reference volumes. However, constraints placed upon the warp are necessary with regards to maintaining a homeomorphic mapping and minimizing metric distortions. As discussed previously, a homeomorphism is biologically necessary for establishing a valid correspondence between homologous structures. To prevent inversions

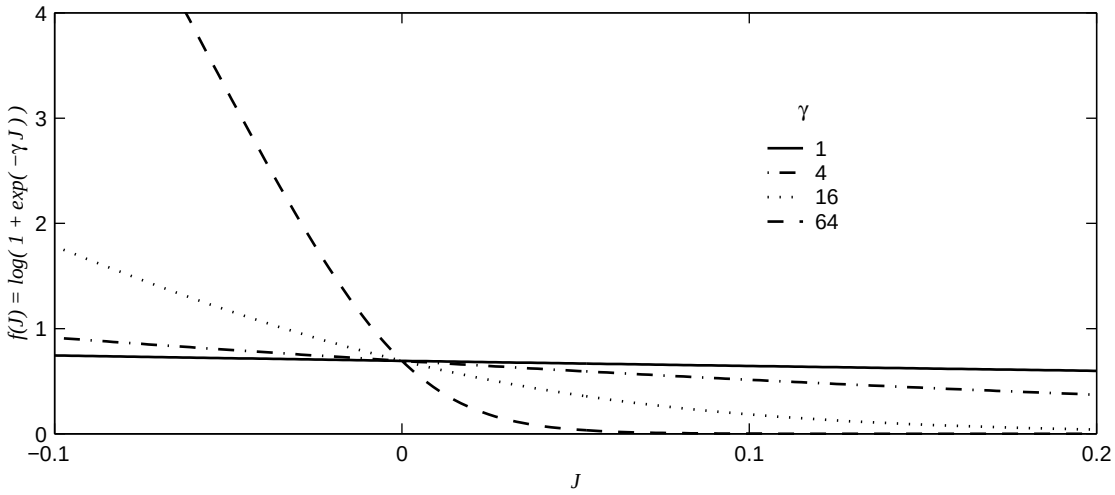

Figure 3.6: *Logarithmic based function of the Jacobian (J) with parameter γ.*

and preserve the homeomorphic property of the mapping, two additional terms are introduced to the objective function.

The second term of the objective function $C_2(\mathbf{D})$ employs the Jacobian matrix of the warp field to ensure that a homeomorphism is produced by the optimization procedure. The Jacobian matrix of a homeomorphic warp field is positive semi-definite at every point. To encourage preservation of this property, $C_2(\mathbf{D})$ is composed of the Jacobian determinant passed through a nonlinear logarithmic based function as shown in Figure 3.6. This function exacts a heavy cost for Jacobian determinants which tend negative, and hence, encourages the optimization to produce homeomorphic warp fields. This term is expressed mathematically as follows:

$$C_2(\mathbf{D}) = \frac{1}{V}\sum_{i=1}^{V} log\left[1+e^{-\gamma J^{(i)}}\right], \tag{3.11}$$

where the determinant of the Jacobian matrix at the i^{th} lattice point of the warp field is denoted as $J^{(i)}$, and as before, V indicates the total number of lattice points over which the warp field is being estimated. The parameter γ controls the severity of the penalty applied to negative Jacobian determinants as is also shown in Figure 3.6.

The Jacobian provides a good measure of the local volumetric distortion induced by the transformation. However, transformations can be constructed that maintain the local volumetric structure and that also induce severe changes in angular shape. (See Subsection 3.1.3 for a geometric interpretation of the Jacobian which further illustrates this issue.) Severe angular shape changes do not make sense from a biologic perspective, and thus the metric distortion term which accounts for both volume and angle is necessary.

The third term of the objective function $C_3(\mathbf{D})$ is designed to minimize volumetric and angular distortion by calculating the amount of relative displacement induced by the warp within some neighborhood and averaged over the entire warp. This calculation is stated explicitly in Equation (3.12).

$$C_3(\mathbf{D}) = \frac{1}{V}\sum_{i=1}^{V}\left[\sum_{n\in N(i)}\left(l^{(in)} - l_0^{(in)}\right)^2\right] \tag{3.12}$$

Where $l^{(in)}$ represents the magnitude of the difference in displacement between the i^{th} sampling position and the n^{th} sampling position as follows:

$$l^{(in)} = \|\mathbf{u}^{(i)} - \mathbf{u}^{(n)}\|_2,$$

and the subscript 0 in $l_0^{(in)}$ indicates the initial relative displacement of these same points as follows:

$$l_0^{(in)} = \|\mathbf{x}^{(i)} - \mathbf{x}^{(n)}\|_2.$$

This term is similar to that introduced by Fischl et al. for minimizing metric distortions associated with a surface-based warp [25].

3.4.4 Gradients

Of course, the minimization of the objective function in a steepest descent fashion requires calculation of the objective function gradient with respect to the set of displacement vectors $\mathbf{D}$. This subsection derives the gradient for each component of the objective function shown earlier in Equation (3.9).

Taking the partial derivative of the intensity term $C_1(\mathbf{D})$ [Equation (3.10)] with respect to the k^{th} displacement vector $\mathbf{d}^{(k)}$ and using $\mathbf{u}^{(k)} = \mathbf{x}^{(k)} + \mathbf{d}^{(k)}$ results in

$$\frac{\partial C_1(\mathbf{D})}{\partial \mathbf{d}^{(k)}} = -\frac{2}{V}\left(\frac{R(\mathbf{x}^{(k)}) - S(\mathbf{u}^{(k)})}{\Omega^2(\mathbf{x}^{(k)})}\right) \cdot \frac{\partial S(\mathbf{u}^{(k)})}{\partial \mathbf{d}^{(k)}}. \tag{3.13}$$

The partial derivative $\frac{\partial S(\mathbf{u}^{(k)})}{\partial \mathbf{d}^{(k)}}$ in the above equation reduces to $\frac{\partial S(\mathbf{u}^{(k)})}{\partial \mathbf{u}^{(k)}}$ by noting the following:

$$\frac{\partial S(\mathbf{u}^{(k)})}{\partial \mathbf{d}^{(k)}} = \frac{\partial S(\mathbf{u}^{(k)})}{\partial \mathbf{u}^{(k)}} \cdot \frac{\partial \mathbf{u}^{(k)}}{\partial \mathbf{d}^{(k)}}, \qquad \frac{\partial \mathbf{u}^{(k)}}{\partial \mathbf{d}^{(k)}} = 1.$$

This allows computation of the gradient based only on the warped intensity function $S(\mathbf{u}^{(k)})$. The gradient at the k^{th} sampling position of the homeomorphic term $C_2(\mathbf{D})$ is computed as

$$\frac{\partial C_2(\mathbf{D})}{\partial \mathbf{d}^{(k)}} = -\frac{1}{V} \cdot \frac{\gamma \frac{\partial J^{(k)}}{\partial \mathbf{d}^{(k)}}}{1 + \exp(\gamma J^{(k)})}. \tag{3.14}$$

Further discussion of how $\frac{\partial J^{(k)}}{\partial \mathbf{d}^{(k)}}$ is calculated can be found in Section 3.5. The gradient of the third term $C_3(\mathbf{D})$ is

$$\frac{\partial C_3(\mathbf{D})}{\partial \mathbf{d}^{(k)}} = \frac{1}{V} \sum_{n \in N(k)} \left[2\left(I^{(kn)} - I_0^{(kn)}\right) \cdot \frac{\partial I^{(kn)}}{\partial \mathbf{d}^{(k)}} \right], \tag{3.15}$$

where $\frac{\partial I^{(kn)}}{\partial \mathbf{d}^{(k)}}$ is a unit vector pointing from the k^{th} sample point to the n^{th} sample point, and $N(k)$ is a neighborhood of the six nearest points about the k^{th} sampling position.

The linear combination of these gradients [Equations (3.13), (3.14), and (3.15)] using λ_1, λ_2, and λ_3 as the respective weighting coefficients provides the gradient used in the steepest descent optimization as shown in the following equation:

$$\frac{\partial C(\mathbf{D})}{\partial \mathbf{d}^{(k)}} = \lambda_1 \frac{\partial C_1(\mathbf{D})}{\partial \mathbf{d}^{(k)}} + \lambda_2 \frac{\partial C_2(\mathbf{D})}{\partial \mathbf{d}^{(k)}} + \lambda_3 \frac{\partial C_3(\mathbf{D})}{\partial \mathbf{d}^{(k)}}.$$

This represents the value of the gradient at the k^{th} sampling position, and the full gradient is found by repeating these calculations at all the sampling points.

3.4.5 Coefficient Selection

To complete the description of the warp optimization a discussion of how the coefficients λ_1, λ_2, and λ_3 are obtained is necessary. Establishing the appropriate relationship between the three terms of the objective function is important for the success of the algorithm, and this relationship is determined by these three coefficients. Actually, one of these coefficients is redundant because it can simply be divided out of the objective function without destroying the relationship between the terms. This observation allows us to set λ_1 equal to unity, and thus, reducing the problem to selection of only two coefficients. The estimation of these remaining coefficients is problematic, however, as no simple, efficient way of determining their optimal values exists. Some work has been done regarding optimal coefficient selection of this type [33], but the computational load of the current algorithm precludes its use.

In this work, these values are obtained heuristically by observing the results of the algorithm. The Jacobian term coefficient (λ_2) is increased if a homeomorphism is not maintained during the generation of the warp field, and the metric distortion coefficient (λ_3) is varied according to a qualitative assessment of the distortion induced by the algorithm.

The final value of the coefficients will vary, however, depending upon the intended application of the resulting warp field. For example, it may be desirable to have a more precise warp at the cost of increased metric distortion. In this case the coefficient for the metric distortion term would be reduced. Other situations certainly exist where the minimization of metric distortion is much more important and hence this coefficient would then be increased.

3.5 Practical Considerations

In order for this algorithm to be executed in a timely manner on general purpose and readily available computational hardware, memory usage and execution speed must be balanced

with the accuracy of the resulting warp. Additionally, the algorithm as stated falls short in some respects necessitating a few somewhat *ad hoc* extensions to the algorithm. These practical considerations are addressed in the following paragraphs.

As noted before, when the algorithm reaches the base level of the Gaussian pyramid where the resolution is the greatest, each vector $\mathbf{d}^{(i)}$ of the warp field $\mathbf{D}$ represents a block of eight voxels. This is done to conserve memory as representation of the full warp field at $256 \times 256 \times 256$ resolution would require approximately 200 MB of computer memory when using single precision values. When combined with the memory required to represent two full Gaussian pyramids (source and reference) which is approximately 150 MB, the limits of readily available processing equipment are stressed. This would unnecessarily extend the processing time of the algorithm by requiring use of very slow virtual memory.

The algorithm as described in the previous sections can have difficulty achieving a smooth warp field, particularly at points close to a transition in intensity levels. To remedy this an additional step is included in the optimization. Before the objective function gradient is used to update the warp field estimate, it is first smoothed by convolving the gradient with a Gaussian shaped blurring kernel (similar to what is done in [25]). At each level of the Gaussian pyramid the degree of smoothing is at first large, and as the optimization converges towards a minimum the amount of smoothing is decreased. This has the effect of encouraging a smooth warp while not adversely affecting the overall match between the reference and the warped source volumes.

One difficulty with steepest descent optimization routines is selection of the step size. An acceptable value can be found through trial and error, but occasionally due to the steep slope of the logarithmic function of the Jacobian (Figure 3.6) a large gradient can be produced which destabilizes the optimization. To control this, the magnitude of the gradient is monitored, and if it exceeds a threshold then the entire gradient is scaled to maintain stability.

Another issue involving the Jacobian is that of computing it based on a discretely sampled warp function. Computation of this quantity could proceed as in Equations (3.3) and (3.4), but there is a problem when computing this in discrete space. In the limit as Δ tends to 0, the differential elements as shown in Figure 3.2 can be accurately represented as parallelopipeds. In the current discrete case, though, Δ represents the distance between sampling points. Here the assumption of parallelopiped shaped volume elements does not hold. All *eight* vertices of the volume element shown in Figure 3.2(a) are transformed somewhat independently and the resulting volume element is in no way constrained to be a parallelopiped. This can lead to inaccurate computation of the Jacobian allowing volumetric inversions to occur unnoticed.

To account for this problem, the true volume of the transformed element should be computed, but this would be computationally intensive. Instead, the following approximation is used. Each vertex of a volume element is also shared by seven other adjacent elements (except for vertices at the outer edges of the full volume). In other words, each vertex is surrounded by eight elements. To calculate the Jacobian determinant at this vertex point, the volumes of the eight surrounding elements are each approximated as parallelopipeds, and after applying the logarithmic based function (Figure 3.6), they are simply averaged together to form the summand of Equation (3.11).

The computation of $\frac{\partial J^{(k)}}{\partial \mathbf{d}^{(k)}}$ in Equation (3.14) is carried out in a similar fashion using the following equation (the superscript (k) has been dropped for simplicity)

$$\begin{aligned}\frac{\partial J}{\partial \mathbf{d}} &= \left(\frac{\partial J}{\partial d_1}\hat{\mathbf{i}}, \quad \frac{\partial J}{\partial d_2}\hat{\mathbf{j}}, \quad \frac{\partial J}{\partial d_3}\hat{\mathbf{k}}\right) \\ &\approx \frac{1}{8}\sum_{\substack{e \in surrounding \\ elements}} \left(\frac{J_{e\pm\Delta\hat{\mathbf{i}}} - J_e}{\Delta}\hat{\mathbf{i}}, \quad \frac{J_{e\pm\Delta\hat{\mathbf{j}}} - J_e}{\Delta}\hat{\mathbf{j}}, \quad \frac{J_{e\pm\Delta\hat{\mathbf{k}}} - J_e}{\Delta}\hat{\mathbf{k}}\right)\end{aligned}$$

where J_e represents the Jacobian determinant computed using the e^{th} surrounding element, and $J_{e\pm\Delta(\hat{\mathbf{i}},\hat{\mathbf{j}},\hat{\mathbf{k}})}$ indicates that this Jacobian is computed with an additional displacement

of $\pm\Delta$ in the $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, or $\hat{\mathbf{k}}$ directions added to the displacement vector $\mathbf{d}$. The additional displacement of $+\Delta$ or $-\Delta$ is consistent with the location of the particular element being utilized. In other words, $+\Delta$ is used for elements located in the $+\hat{\mathbf{i}}$, $+\hat{\mathbf{j}}$, and $+\hat{\mathbf{k}}$ directions and $-\Delta$ is used for elements located in the $-\hat{\mathbf{i}}$, $-\hat{\mathbf{j}}$, and $-\hat{\mathbf{k}}$ directions.

A final point of minor consequence is that the termination of the optimization is based solely upon the similarity measure $C_1(\mathbf{D})$ [Equation (3.10)] and not the value of the objective function as a whole [Equation (3.9)]. Once the similarity measure asymptotes or begins to increase signaling that no more improvement is possible, the optimization at that level of the Gaussian pyramid concludes and the algorithm proceeds to the next level. By monitoring the progress of the optimization in this way the regularization terms are allowed to become as large as necessary so as to maintain a homeomorphism without significantly affecting the optimization.

Chapter 4

Results

Details concerning the testing and analysis of the results are presented in this chapter for the current algorithm. Section 4.1 provides an overview of the volumetric image datasets used for testing, and Section 4.2 discusses the performance.

The overall goals of an automatic, precise, robust, and objective algorithm have only partially been achieved. The algorithm is automatic in that it requires no human intervention once it has been initiated, and the algorithm is objective in that it is deterministic and will not change its result over repeated executions on the same data. However, sufficient levels of precision and robustness have not been achieved. Nonetheless, there are some interesting results regarding the performance of the algorithm.

4.1 Volumetric Image Set Overview

4.1.1 Data Calibration

High resolution T1-weighted images generated by an MR scanner such as those used in this work are typically corrupted by magnetic field and radio frequency inhomogeneities. This results in variations in intensity and contrast across the image. That is, identical tissue

types (e.g. cortical gray and white matter) will give rise to varying intensities as a function of their spatial location. This is obviously undesirable for any registration procedure which utilizes intensity information in order to match tissue types between the reference and source volumes. Typically the variation is most pronounced in the *z* (superior-inferior) direction. In order to correct for this, a procedure for computing the one-dimensional "bias field" and applying its inverse to the image to generate normalized intensity values is used. The procedure relies on the assumption that in any given slice parallel to the *x-y* plane in the magnet coordinate system, the highest intensity tissue type of any significance will be white matter. The algorithm proceeds as follows:

1. Construct a set of histograms from overlapping slices parallel to the *x-y* plane in the magnet coordinate system. On the order of 10-15 slices with 50% overlap in adjacent slices is typically used.

2. Smooth the resulting histograms using a fairly broad ($\sigma = 2$) Gaussian to eliminate noise-induced peaks.

3. Use a peak-finding algorithm to determine the rightmost (i.e. highest intensity) peak in the histogram containing a significant number of voxels (usually more than 15% of the non-background voxels in the slice). The location of the peak specifies the mean white matter intensity for that slice.

4. Discard outliers from the array of detected mean white matter intensities. This step makes use of the fact that the variation in intensity due to magnetic field inhomogeneities is smooth across space, and that therefore the variation in detected mean white matter intensity should be small in any two adjacent slices. Specifically, the median detected white matter value in a set of slices around the Talairach origin is used as a starting point. Then slices are added for which the gradient of the intensity change over space with respect to the nearest valid slice is below a prespecified threshold (0.4/mm).

5. Fit a set of cubic splines to the resulting coefficients of the valid slices.

6. Use the splines to interpolate the coefficients for each point along the *z*-axis.

7. Adjust each intensity value by the coefficient at its *z* coordinate.

This procedure is quite robust due to the minimal assumptions on which it relies (i.e. no parametric form) as well as the built-in cross-validation in step (4).

4.1.2 Source Volumes

The volumetric image sets utilized for this work are T1-weighted volumes with $1\times1\times1$ mm voxel resolution taken using a short 3-D acquisition sequence (eg. GE SPGR, Siemens MP-RAGE). The dimensions of the volume covering the brain are approximately 16.9 cm laterally, 25.6 cm inferior to superior, and 25.6 cm posterior to anterior. The resulting data volume is $256\times256\times256$ voxels where the lateral dimension has been padded with zeros.

Two normal volumes have been selected to facilitate performance analysis of the algorithm. Figure 4.1 shows the sagittal, coronal, and transverse slices of Volume A. Figure 4.2 shows the corresponding slices of Volume B.

4.1.3 Reference Volumes

The reference volume that is used for registration is constructed by averaging N previously aligned individual brain volumes. This is expressed in the following equation:

$$R(\mathbf{x}) = \frac{1}{N}\sum_{k=1}^{N} S_k(\mathbf{x}),$$

where S_k represents the k^{th} aligned volume. Associated with this average volume is a measure of the intensity variance at each voxel location ($\Omega^2(\mathbf{x})$) which is estimated in an unbiased fashion as follows:

$$\Omega^2(\mathbf{x}) = \frac{1}{N-1}\sum_{k=1}^{N}\left[S_k(\mathbf{x}) - R(\mathbf{x})\right]^2.$$

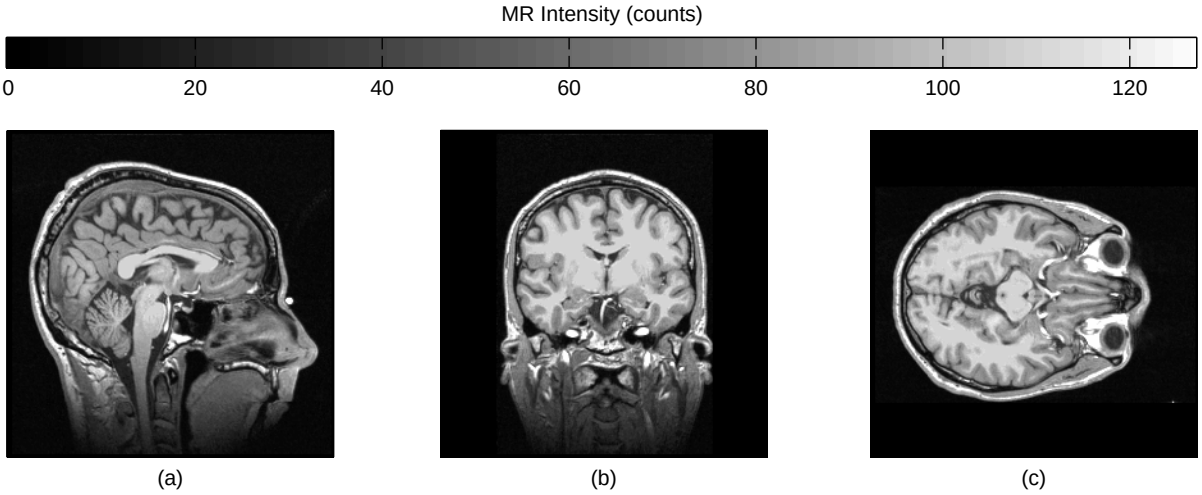

Figure 4.1: *Image volume of a typical individual brain (Volume A): (a) sagittal, (b) coronal, and (c) transverse slices.*

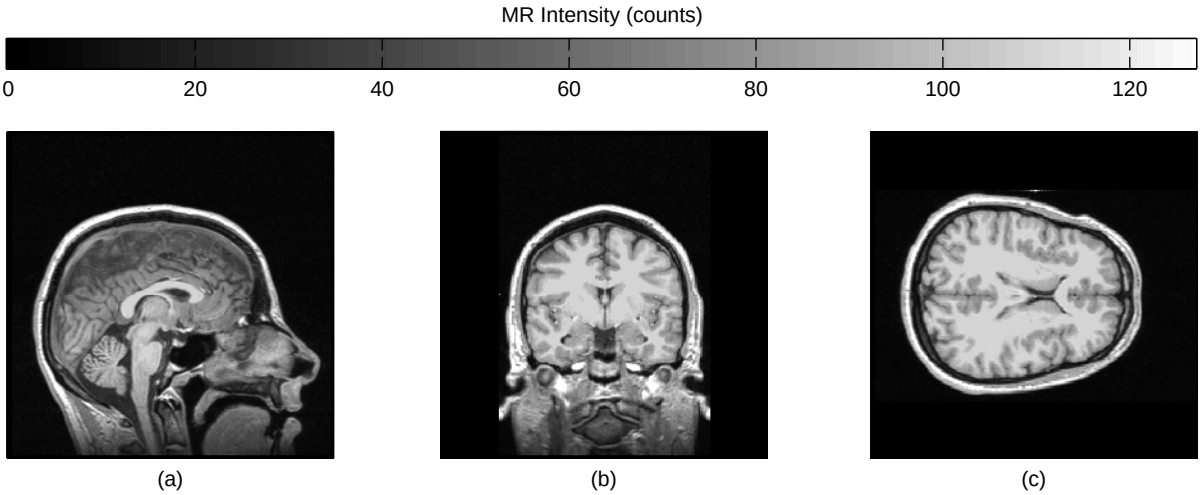

Figure 4.2: *Image volume of a second typical individual brain (Volume B): (a) sagittal, (b) coronal, and (c) transverse slices.*

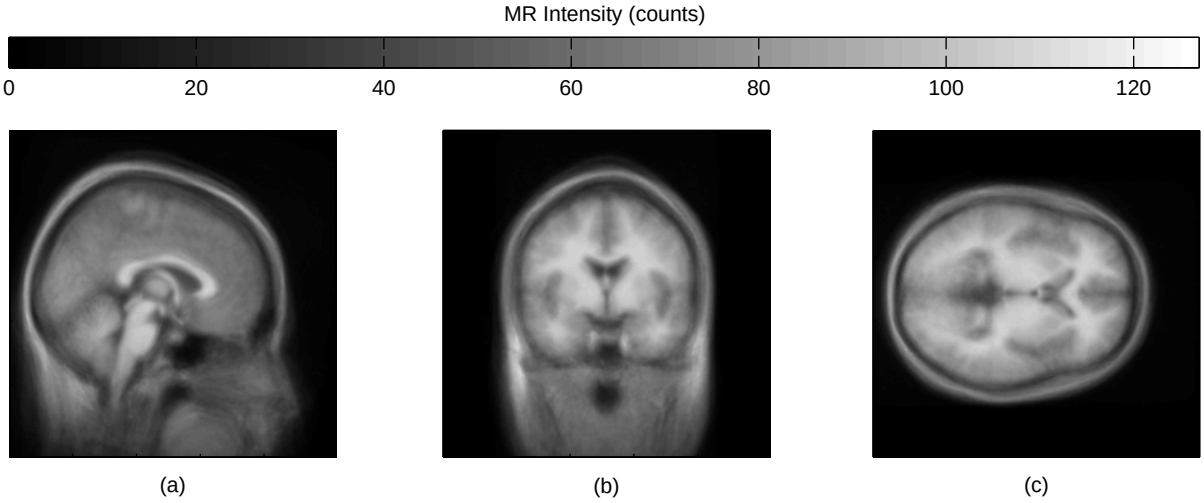

Figure 4.3: *Reference data volume constructed by averaging 40 Talairach registered brains: (a) sagittal, (b) coronal, and (c) transverse slices.*

For the initial reference volume used in this work, 40 volumes are registered in the Talairach registration space and then averaged as above. The resulting composite volume is shown in Figure 4.3, and corresponding slices through the $\Omega(\mathbf{x})$ data volume are shown in Figure 4.4.

Visual inspection of this initial reference volume (Figure 4.3) and comparison with a typical individual volume (Figure 4.1) seems to show that the intrinsic resolution of the reference volume is less than the $256\times256\times256$ voxels used to represent it. A crude qualitative comparison between the sagittal slice of the reference [Figure 4.3(a)] and a Gaussian pyramid constructed from an individual volume (Figure 4.5) indicates that the true intrinsic resolution varies somewhere between $32\times32\times32$ voxels and $128\times128\times128$ voxels. The corpus collosum and other subcortical features seem to correspond well with the $128\times128\times128$ voxel representation [Figure 4.5(a)] while the cortical gyri and sulci seem to only be resolved as well as those of the $32\times32\times32$ voxel resolution individual [Figure 4.5(c)].

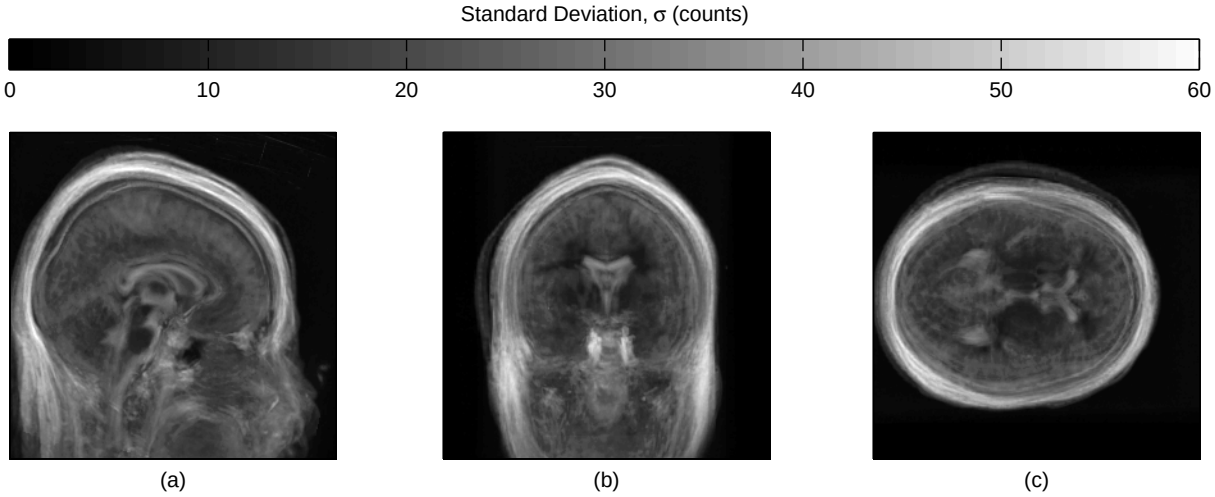

Figure 4.4: *Standard deviation values (Ω) associated with the reference volume (Figure 4.3): (a) sagittal, (b) coronal, and (c) transverse slices.*

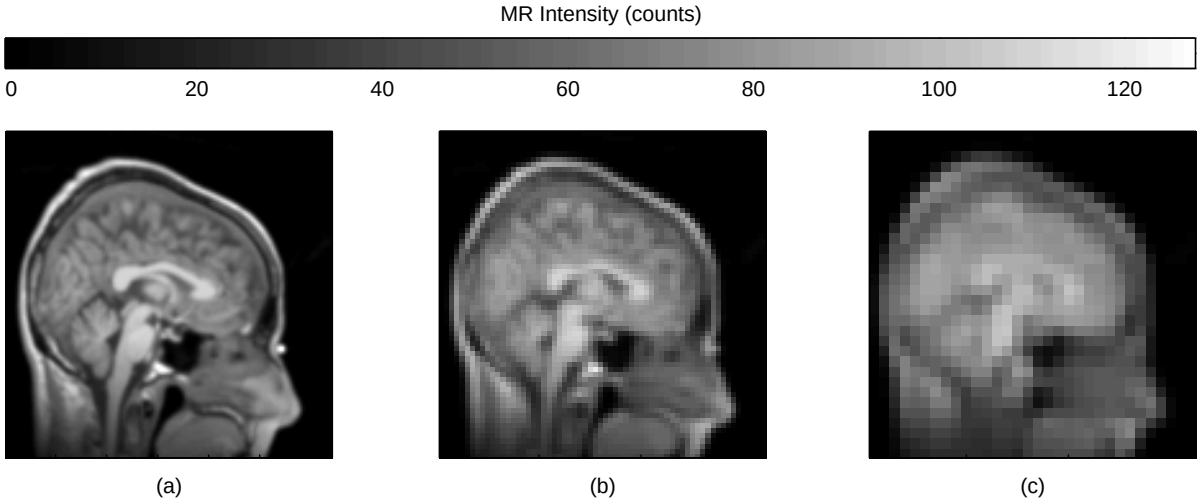

Figure 4.5: *A mid-sagittal slice of a brain volume represented at three levels of a Gaussian pyramid: 128×128×128 (a), 64×64×64 (b), and 32×32×32 (c).*

This observation raises some issues as to why this composite volume with an intrinsically lower resolution is used as a reference for a high resolution individual. Ultimately the goal is to not use this particular reference volume, but to use it as an intermediate step which will help to obtain warps which improve upon the Talairach registrations. The same volumes used to generate the reference can be realigned with the current algorithm and then a second reference can be constructed which will presumably have a higher intrinsic resolution and less variability than the current reference volume. This step can be repeated as appropriate to obtain a reference with a suitably high resolution. Due to time constraints, however, only the initial reference volume is utilized in this work.

4.2 Algorithm Performance

4.2.1 Error Analysis

To evaluate the performance of the algorithm, the two individual volumes (A and B) are first warped to the composite reference volume, and comparisons are then made between the two resulting warped volumes. The objective function coefficients λ_1, λ_2, and λ_3 are set to 1, 1×10^{-3}, and 1×10^{-7} respectively, and the primary metric utilized for comparison is the mean squared error (MSE). Table 4.1 summarizes the MSE between Volume A and Volume B at various stages of the algorithm and at various resolutions.

Table 4.1: Mean squared error (MSE) between Volume A and Volume B.

	Pyramid Level			
Warp Field	32×32×32	64×64×64	128×128×128	256×256×256
Affine	61.46	134.98	226.74	**367.82**
32×32×32	**20.45**	86.59	188.16	339.00
64×64×64	24.37	**69.22**	164.53	308.63
128×128×128	24.38	75.41	**164.54**	308.44

Bold face indicates the pyramid level actually utilized to generate the particular warp field.

To further explain Table 4.1, the warping algorithm is executed on each source volume

(A and B) independently using the previously described reference volume in both cases. In the course of the algorithm, first an affine warp field is constructed and then warp fields of $32\times32\times32$, $64\times64\times64$, and $128\times128\times128$ resolution are successively generated. Each warp field is actually obtained utilizing only a single level of the Gaussian pyramid (the **bold face** entries in the table), but then is applied to all the levels for the purposes of analysis. To accomplish this, it is sometimes necessary to apply a lower resolution warp field to a higher resolution volume. Conversely, a higher resolution warp field is, at times, applied to a lower resolution volume. For this analysis, the appropriately sized warp field is generated by using trilinear interpolation. Now, applying the warp fields to the volumes generates a set of 16 different warped volumes (four warp fields $\times$ four pyramid levels) for each source volume. The table entries indicate the MSE between corresponding warped volumes of Volume A and Volume B.

Reading across the table from left to right, the MSE always increases. This can be understood by realizing that at each lower level (higher resolution) of the Gaussian pyramid more and sharper detail is revealed, thus increasing the potential for mismatched intensities. Equivalently, the MSE always decreases when reading in the opposite direction (right to left) because the volumes being compared are successively blurred and subsampled. This blurring and subsampling has the effect of making the volumes more similar.

A more interesting view of Table 4.1 is to read from top to bottom. Here we see that the MSE initially improves when advancing from the affine warp field to the $32\times32\times32$ warp field. The $64\times64\times64$ warp field, however, does not always improve the match (the $32\times32\times32$ level of the pyramid becomes worse going from 20.45 MSE to 24.37 MSE). Furthermore, progressing to the $128\times128\times128$ warp field we see essentially no improvement at any level of the pyramid. This behavior can be explained by recalling that the intrinsic resolution of the reference volume is close to $64\times64\times64$ as described in Subsection 4.1.3. The warping algorithm cannot be expected to achieve improved results at levels beyond the intrinsic resolution of the reference volume being employed.

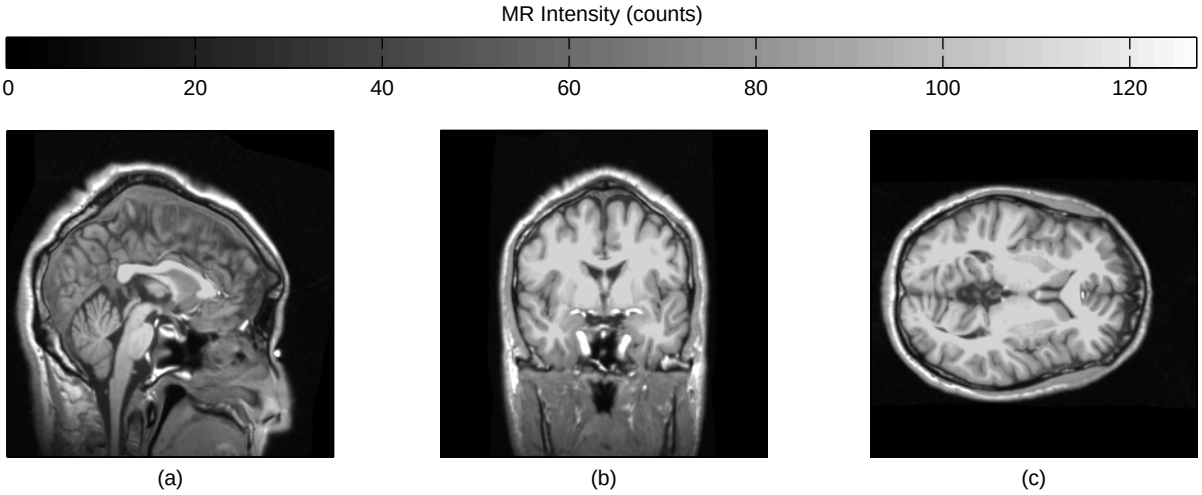

Figure 4.6: *Warp of Volume A using the 128×128×128 derived warp field: (a) sagittal, (b) coronal, and (c) transverse slices.*

Presumably, as discussed in Subsection 4.1.3, the warped volumes that are produced by using this particular reference volume will be more suitably aligned, thus enabling a reference volume with higher intrinsic resolution to be constructed. Interpretation of the results that follow do not necessarily support this presumption, however. Figure 4.6 shows Volume A after application of the 128×128×128 warp field. Artifacts have been introduced in attempting to match beyond the intrinsic resolution of the reference volume. For example, the wavy quality of the scalp and the corpus collosum as seen in the sagittal view of Figure 4.6 does not look particularly natural. Figure 4.7 shows the effect of the independently derived 128×128×128 warp field applied to Volume B. Other unnatural artifacts appear here such as the shadow-like blurring of the scalp on the left and right sides of the coronal slice and the appearance of high intensity dots outside of the scalp in the upper right of the transverse view.

The difference map for these two warped volumes is shown in Figure 4.8. For comparison we show the affine registered versions of Volume A and Volume B in Figures 4.9 and 4.10 respectively along with their difference map in Figure 4.11.

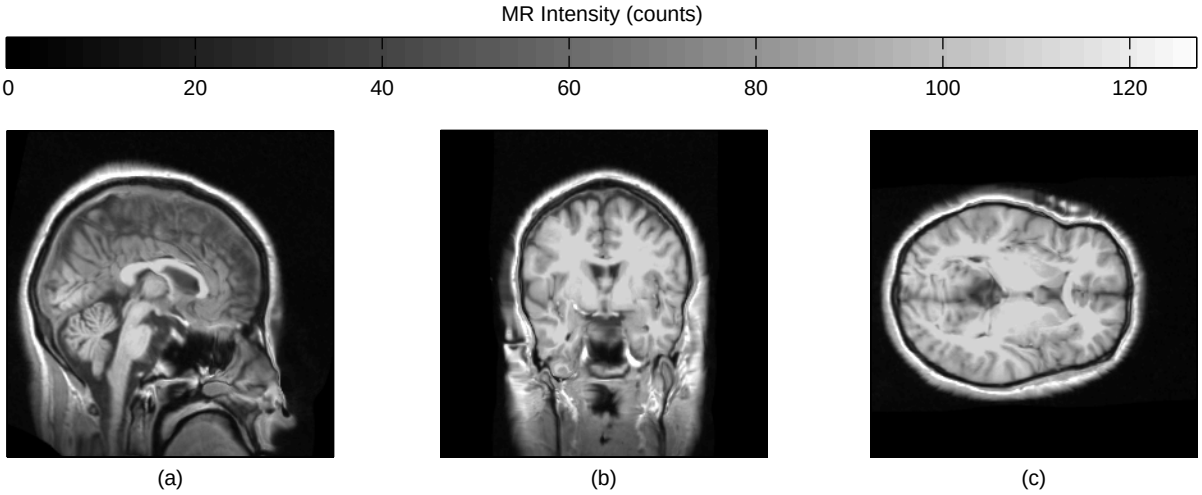

Figure 4.7: *Warp of Volume B using the 128×128×128 derived warp field: (a) sagittal, (b) coronal, and (c) transverse slices.*

In terms of the MSE, the errors between Volume A and Volume B when warped with their respective 128×128×128 warp fields are smaller than the errors produced with their affine transformations. The difference in MSE is only 16%, however. Note that the affine transformation appears qualitatively superior as none of the previously mentioned artifacts are present.

At a resolution of 32×32×32, the visual quality of the warped volumes is improved somewhat as seen in Figure 4.12. This represents a 60% improvement in the MSE over the affine transformation without introducing any significant artifacts. For comparison the affine transformation at this same level of resolution is shown in Figure 4.13.

4.2.2 Constraint Analysis

Minimizing the MSE is only one aspect of achieving a suitable warp. As described before, additional constraints are placed upon the algorithm to facilitate a biologically reasonable warp. These constraints are designed to maintain a homeomorphism and also to minimize the amount of distortion.

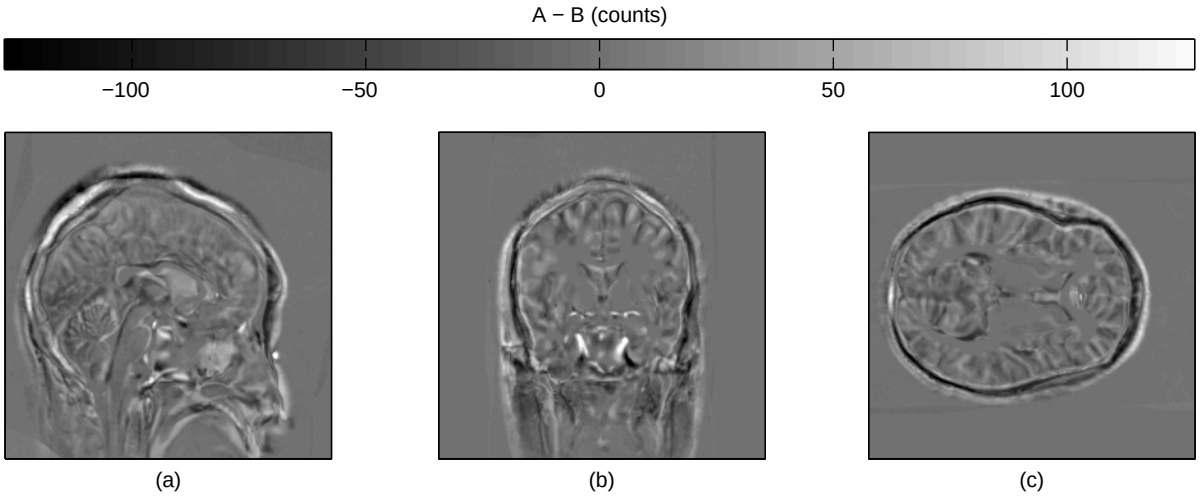

Figure 4.8: *Difference map of Volumes A and B warped with the 128×128×128 derived warp field (Figures 4.6 and 4.7 respectively): (a) sagittal, (b) coronal, and (c) transverse slices.*

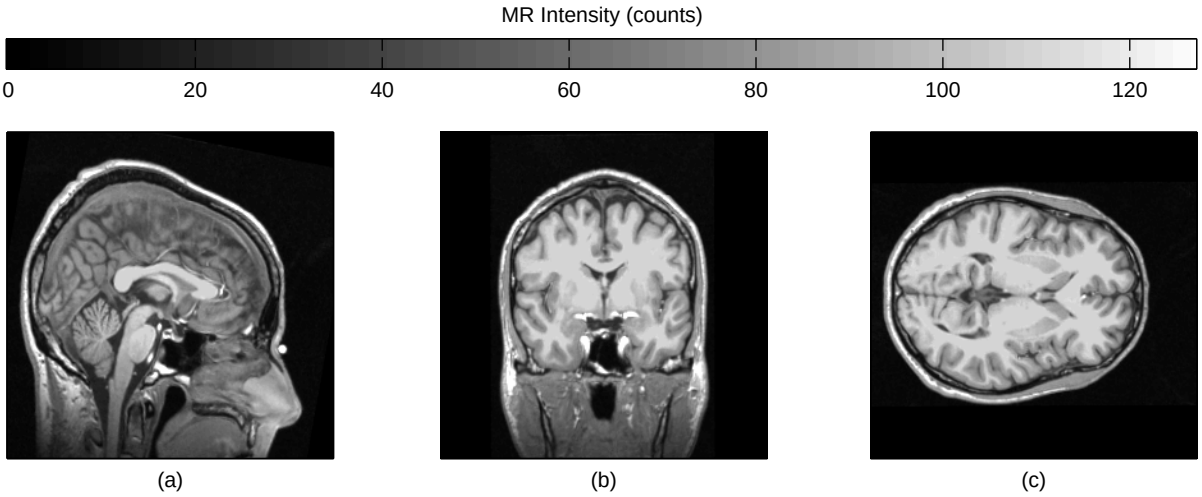

Figure 4.9: *Affine registration of Volume A with the reference volume: (a) sagittal, (b) coronal, and (c) transverse slices.*

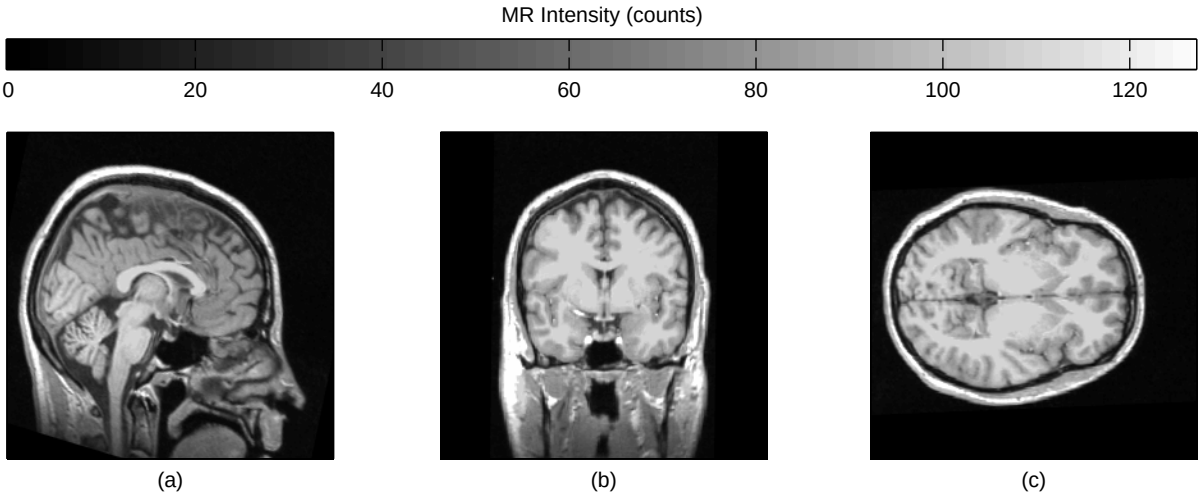

Figure 4.10: *Affine registration of Volume B with the reference volume: (a) sagittal, (b) coronal, and (c) transverse slices.*

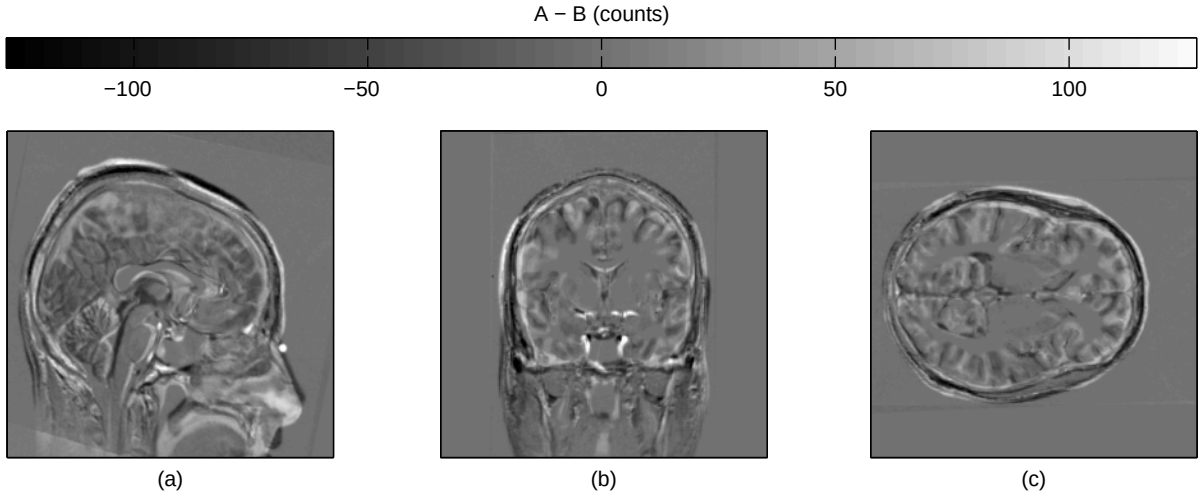

Figure 4.11: *Difference map of Volumes A and B after application of the affine warp: (a) sagittal, (b) coronal, and (c) transverse slices.*

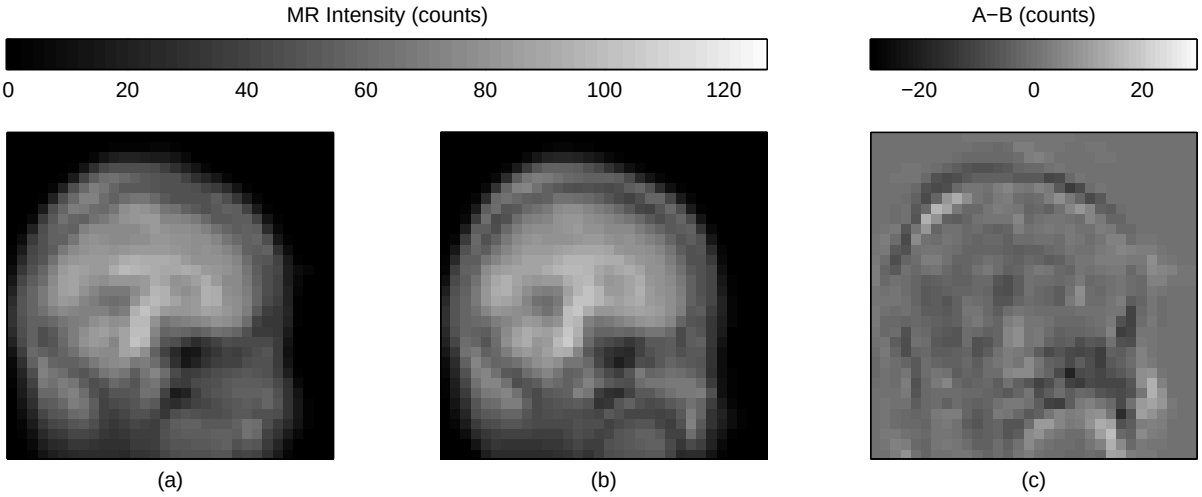

Figure 4.12: *Sagittal view after nonlinear warping of (a) Volume A, (b) Volume B, and (c) their difference (A-B) at 32×32×32 resolution.*

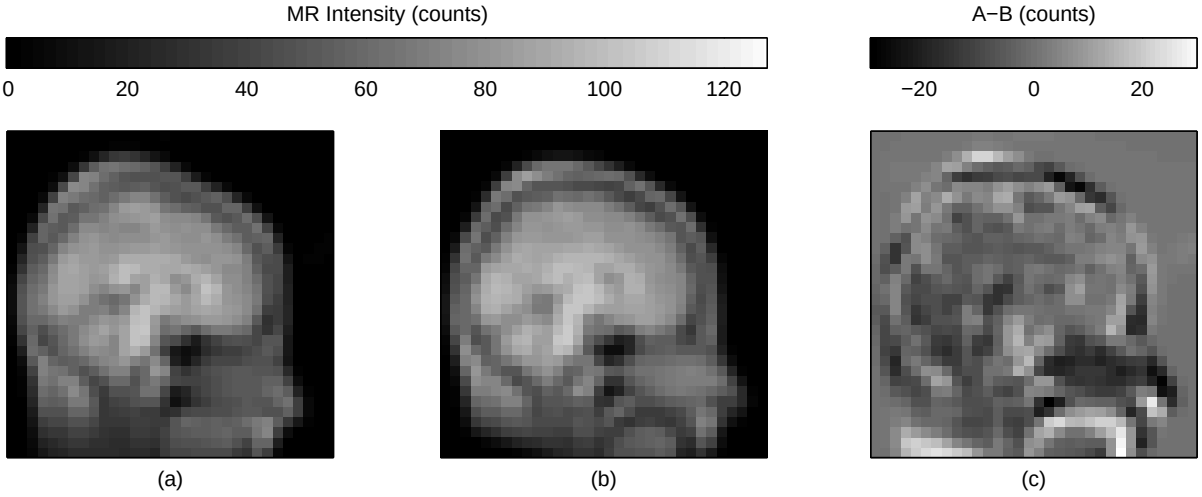

Figure 4.13: *Sagittal view after affine transformation of (a) Volume A, (b) Volume B, and (c) their difference (A-B) at 32×32×32 resolution.*

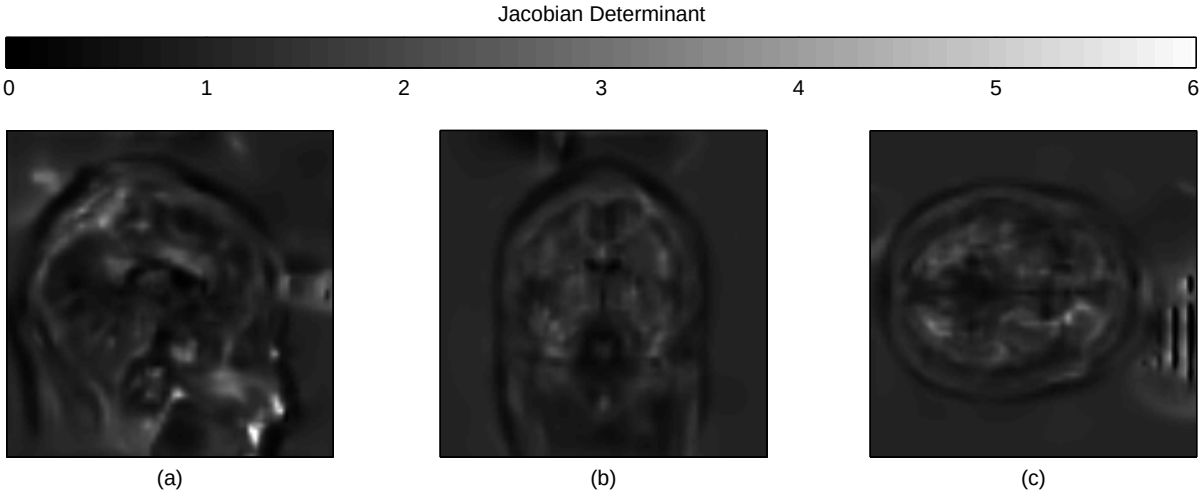

Figure 4.14: *Jacobian map for the warp field obtained for Volume A: (a) sagittal, (b) coronal, and (c) transverse slices.*

Shown in Figure 4.14 is a map of the Jacobian determinant for the final warp field obtained from the algorithm after application to Volume A. The three views of this figure correspond exactly to the views shown in Figure 4.9. The views of the Jacobian determinant are dominated by values close to unity. Thus the map appears rather dark which indicates that most of the volume is being preserved. In addition, there are some points of higher intensity indicating volumetric expansion is occurring. On the right side of the sagittal view of Figure 4.14, there exist a few points that are less than zero indicating that volumetric inversion has occurred. These points, however, are less consequential since they are outside of the brain volume itself.

Figure 4.15 shows the metric distortion induced by the algorithm. Here we see that metric distortion is dominated by small values with only a few locations exceeding more than 2 mm. The area on the right side of the sagittal view does exhibit relatively large amounts of distortion, but again, this area is outside of the brain volume itself. The explanation, however, for the relatively large distortion and volumetric inversion is unknown.

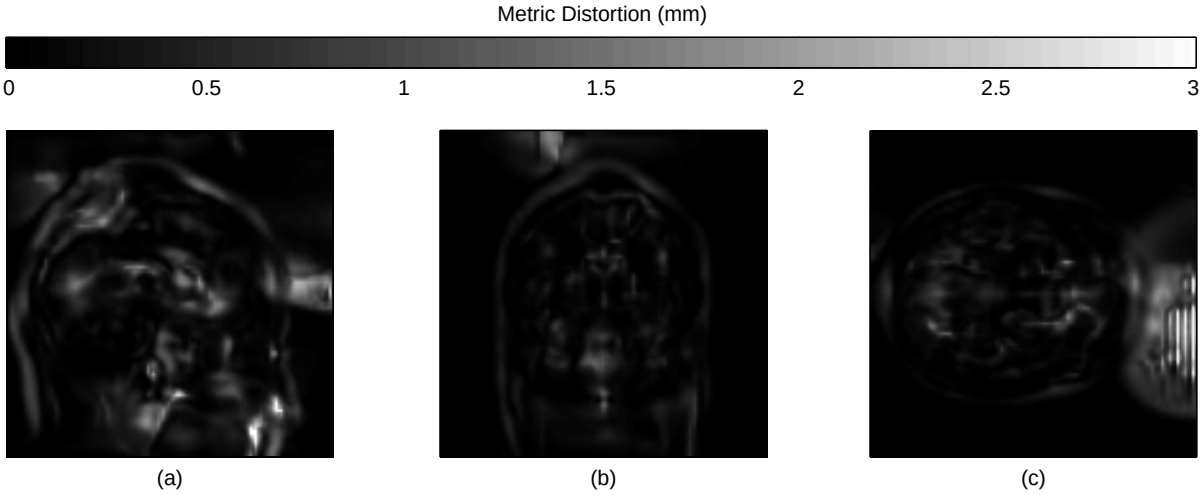

Figure 4.15: *Metric distortion introduced by the warp field obtained for Volume A: (a) sagittal, (b) coronal, and (c) transverse slices.*

4.2.3 Execution Speed

The execution of the algorithm can take a considerable amount of time. For the typical source and reference volumes shown in the previous sections, the algorithm requires approximately 80 hours of execution time on a 700 MHz Pentium III system with 512 MB of main memory. A large amount of this time is spent in easily programmed but inefficient loops as well as recording intermediate diagnostic results to aid in the development of the algorithm. No attempt was made to improve this aspect of performance, however, and certainly reduced execution times could be obtained through optimized programming techniques and removal of the diagnostic code.

Chapter 5

Conclusions

As noted previously, the current technique does not completely satisfy the requirements for a suitable warp, so discussion of potential modifications is provided in Section 5.1. Additionally, thoughts on the validity of brain warping in general are contained in Section 5.2, and concluding remarks are found in Section 5.3.

5.1 Possible Modifications

Reduced execution time would greatly facilitate the further development of the current algorithm. While overall speed of the algorithm is not of particular technical interest here, the exceedingly long execution times of 80 hours hinder the efficient evaluation of algorithm modifications and extensions. Software profiling can be utilized to identify the inefficient portions of the algorithm code which can then be optimized in some fashion. Additionally, once development of the algorithm has been completed, the production version of the code can benefit from removal of the intermediate diagnostic output.

When reasonable execution times are achieved, a more thorough and systematic search for suitable objective function coefficients (λ_1, λ_2, and λ_3) can be accomplished. In terms

of MSE and constraint performance, finding more suitable coefficients is probably the simplest and most direct way to improve the algorithm.

Another aspect that should be revisited is the similarity term of the objective function [Equation (3.10)]. This term is constructed to help guide the generation of the warp by minimizing a modified mean squared error. The squared difference in this term is modulated by an estimate of the variance at each sampling location (Ω^2). The motivation for this is that highly variable locations are less reliable and hence should contribute less to the evolution of the warp field. As can be seen in Figure 4.4, the area that is the most highly variable is the skull and scalp, but this is precisely the region that has the poorest match after warping Volume A and Volume B as shown in Figure 4.8. Clearly, this region must be more vigorously and accurately aligned which suggests removing the Ω^2 term. Removal of this term, though, may be problematic. The skull and scalp in the reference volume shown in Figure 4.3 are not clearly delineated and are of a much lower intensity than the skull and scalp of the source volumes in Figures 4.1 and 4.2. (This blurred quality of the skull and scalp region is, of course, due to the higher variability of this region.) Consequently, the reference volume does not provide much guidance for the warping algorithm as to how the skull and scalp of the source volumes should be registered. The solution for this problem is unclear, but potentially a high resolution individual who is somehow representative of normal brains should be utilized as the reference volume. This would allow removal of the Ω^2 term in the objective function, and the high resolution reference would provide improved guidance for the warping algorithm.

5.2 Warp Validity

Brain warping is an active area of research which combines both mathematics and signal processing with biology and neuroanatomy. Considerable attention has been paid to the mathematical and signal processing aspects as evidenced by the wide variety of techniques

which rely on esoteric mathematical constructs such as Christoffel symbols [34], procrustes metrics [35], continuum-mechanics [9], etc. While these mathematical formalisms are interesting, the vast majority of warping techniques emphasize the mathematics and neglect the critical component of biology. Without a strong connection to the physical and biological properties of living tissues, warping algorithms degenerate into a somewhat arbitrary realignment of voxels. Studies of anatomy or brain activation based on these arbitrary transformations may not contain significant meaning.

To be meaningful, warping algorithms should be strongly rooted in a biological framework which incorporates not only reasonable biological constraints but also the natural variation of the anatomy under investigation. Ashburner and Friston come the closest to employing biologically relevant constraints by suggesting the use of empirically derived priors for their MAP based scheme [12]. Since the macroscopic physical nature of brain tissue is relatively unknown, utilization of data based estimates of variance seems particularly compelling. Even though the physical properties are unknown or difficult to model, the warp will still progress in ways that are similar to the naturally occurring variations of the brain.

Most warping algorithms constrain the warp in some fashion such that obviously incorrect warps which violate basic biological principles are avoided. For example, it is generally accepted in the literature that smooth homeomorphic warps are required for the warp to be considered biologically adequate. However, the degree of smoothness required is an open question. Even when a vague answer is given it is usually application specific. Additionally, homeomorphic maps are sought at the expense of matching anatomic features. While the gross anatomical structures have specific relationships which require a homeomorphism at large scales, the matching of details at much finer scales may be unnecessarily blocked by the desire to maintain a pure homeomorphism. The absence of obvious biologic incompatibilities does not preclude the existence of non-obvious biologic problems which should be sought out and addressed.

The concise specification of biologically relevant constraints with regard to brain warping is a difficult and unsolved problem, and even once these are clearly stated in biological terms, their translation into mathematical constructs can be particularly problematic. Nonetheless, the implications of biological validity should still be considered. Currently available warping algorithms do have applicability in clinical and research settings, but it is vitally important that the warping methods employed are well understood with regard to their biological implications. The application of the warp must be considered when selecting a warping algorithm, and if the algorithm does not provide sufficient meaning in the chosen context then an alternative algorithm should be found or constructed.

The current selection of algorithms seems to be most suited to situations where a direct one-to-one mapping is known to exist as in the case of registration of a single brain across imaging modalities or across time. In these cases, a single specific brain is under consideration, and unless some physical trauma has occurred between scans, there will be a direct homeomorphic correspondence between the image volumes.

5.3 Concluding Remarks

This report began with a brief history and background of brain warping, as well as discussing the wide variety of applications. Several of the more prominent warping algorithms were also discussed with respect to the current algorithm. These included the Talairach atlas which is widely used due to its simplistic implementation, intensity based algorithms which rest primarily on the image intensities, and model based algorithms which rely upon extraction of geometric features to effect the warp. A full discussion of the mathematics behind transformations is then provided and used as the basis for the development of the current algorithm.

Fundamentally an intensity based algorithm, the technique seeks to minimize an objective function which is constructed to encourage similarity in intensity while discouraging

volumetric inversion and metric distortion. This is accomplished in a robust fashion by utilizing a multi-resolution scheme. Furthermore, the algorithm operates with no human intervention once it is initiated. Results are then analyzed in terms of the mean squared error (MSE) as well as with regard to the constraints incorporated into the algorithm.

The largest unsolved problem with the current algorithm is the appropriate selection of the objective function coefficients (λ_1, λ_2, and λ_3). The relationship between the objective function terms as specified by these coefficients is vital to the success of the algorithm, but the techniques available in the literature for obtaining them would require an intractable amount of computation if applied to the current algorithm [33]. Also contained in the concluding chapter is a discussion of the general validity of warping algorithms and a suggestion that the internal workings of any warping technique must be understood with respect to the biological nature of brains before application to a specific problem.

Bibliography

[1] George H. W. Bush, Presidential Proclamation 6158, July 17, 1990.

[2] M. I. Sereno, J. M. Allman, “Coritical visual areas in mammals,” in *The Neural Basis of Visual Function*, A.G. Leventhal, Ed., Macmillan, London, 1991, pp. 160-172.

[3] D. Felleman and D. C. Van Essen, “Distributed hierarchical processing in primate cerebral cortex,” *Cerebral Cortex*, Vol. 1, 1991, pp. 1-47.

[4] A. Morel, J. H. Kaas, “Subdivisions and connections of auditory cortex in owl monkeys,” *Journal of Comparative Neurology*, Vol. 318, 1992, pp. 27-63.

[5] M. M. Merzenich and J. F. Brugge, “Representation of the cochlear partition on the superior temporal plane of the macaque monkey,” *Brain Research*, Vol. 50, 1973, pp. 275-296.

[6] M. M. Merzenich, J. H. Kaas, M. Sur, C.-S. Lin, “Double representation of the body surface within cytoarchitectonic areas 3b and 1 in “S-I” in the owl monkey,” *Journal of Comparative Neurology*, Vol. 181, 1978, pp. 41-74.

[7] I. Stepniewska, T. M. Preuss, and J. H. Kaas, “Architectonics, somatotopic organization, and ipsilateral cortial connections of the primary motor area (M1) of owl monkeys,” *Journal of Comparative Neurology*, Vol. 330, 1993, pp. 238-271.

[8] Jean Talairach and Pierre Tournoux, *Co-planar stereotaxic atlas of the human brain*, Mark Rayport Translator, Thieme Medical Publishers, New York, 1988.

[9] Arthur W. Toga, Editor, *Brain Warping*, Academic Press, San Diego, 1999.

[10] Commission on Physical Sciences, Mathematics, and Applications, *Mathematics and Physics of Emerging Biomedical Imaging*, National Academy Press, Washington, D. C., 1996.

[11] Paul M. Thompson, et al., "Growth patterns in the developing brain detected by using continuum mechanical tensor maps," *Nature*, Vol. 404, 9 March 2000.

[12] John Ashburner and Karl J. Friston, "Chapter 2: Spatial Normalization," *Brain Warping*, Arthur W. Toga, Ed., Academic Press, San Diego, 1999.

[13] J. Talairach, M. David, P. Tournoux, H. Corredor, and T. Kvasina, *Atlas d'anatomie stereotaxique des noyaux gris centraux*. Masson, Paris, 1957.

[14] J. Talairach, G. Szikla, P. Tournoux, A. Prossalentis, M. Bordas-Ferrer, L. Covello, M. Jacob, A. Mempel, P. Busser, and J. Bancaud, *Atlas d'anatomie stereotaxique telencephale*. Masson, Paris, 1957.

[15] J. Hunter Downs III, Jack L. Lancaster, and Peter T. Fox, "Chapter 15: Surface-Based Spatial Normalization Using Convex Hulls," *Brain Warping*, Arthur W. Toga, Ed., Academic Press, San Diego, 1999.

[16] John W. Haller, "Chapter 17: Brain Templates," *Brain Warping*, Arthur W. Toga, Ed., Academic Press, San Diego, 1999.

[17] D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans, "Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space," *Journal of Computer Assisted Tomography*, Vol. 18, pages 192-205.

[18] D. C. Van Essen, and H .A. Drury, *Structural and Functional Analysis of Human Cerebral Cortex Using a Surface-Based Atlas*, The Journal of Neuroscience, Vol. 17, No. 18, 1997, pp. 7079-7102.

[19] H. Steinmetz, G. Furst, and H. J. Freund, *Cerebral Cortical Localization: Application and Validation of the Proportional Grid System in MR Imaging*, J. Comput. Assist. Tomography, Vol. 13, pp. 10-19.

[20] D. L. Hunton, F. J. Miezen, R. L. Buckner, H. I. Mier, M. E. Raichle, and S. E. Petersen, *An assessment of functional-anatomical variability in neuroimaging studies*, Human Brain Mapping, Vol. 4, 1996, pp. 122-139.

[21] P. M. Thompson, and A. W. Toga, *A surface-based technique for warping 3-dimensional images of the brain*, IEEE Transactions on Medical Imaging, Vol. 15, 1996, pp. 1-16.

[22] Louis Collins and A. C. Evans, "Chapter 8: ANIMAL: automatic nonlinear image matching and anatomical labeling," *Brain Warping*, Arthur W. Toga, Ed., Academic Press, San Diego, 1999.

[23] J. Ashburner and K. Friston, "Fully three-dimensional nonlinear spatial normalisation: a new aproach," *2nd International Converence on Functional Mapping of the Human Brain*, Organization for Human Brain Mapping, Boston, 1996.

[24] A. M. Dale, B. Fischl, and M. I. Sereno, "Cortical Surface-Based Analysis I: Segmentation and Surface Reconstruction," NeuroImage 9(2):179-194, 1999.

[25] Bruce Fischl, M. I. Sereno, Anders M. Dale, "Cortical Surface-Based Analysis II: Inflation, Flattening, and a Surface-Based Coordinate System," NeuroImage 9(2):195-207, 1999.

[26] J. C. Gee, D. R. Haynor, L. Le Briquer, R. K. Bajcsy, "Advances in elastic matching theory and its implementation," *CVRMed-MRCAS'97*, Eds. G. Goos, J. Hartmanis, J. Van Leeuwen, Spring-Verlag, New York, 1997.

[27] P. M. Thompson and A. W. Toga, "Visualization nad mapping of anatomic abnormalities using a probabilistic brain atlas based on random fluid transformations," *Proceedings of the Interantional Conference on Visualization in Biomedical Computing*, K. H. Hohne, and R. Kikinis, Eds., Springer-Verlag, New York, 1996, pp. 383-392.

[28] Gérard Subsol, "Chapter 14: Crest lines for curve-based warping," *Brain Warping*, Arthur W. Toga, Ed., Academic Press, San Diego, 1999.

[29] Manfredo P. do Carmo, *Differential geometry of curves and surfaces*, Prentice-Hall, Inc., New Jersey, 1976.

[30] Richard A. Hunt, *Calculus with Analytic Geometry*, Harper & Row Publishers, New York, 1988.

[31] Paul S. Heckbert, "Fundamentals of Texture Mapping and Image Warping," Master's Thesis, (Technical Report No. UCB/CSD 89/516), University of California, Berkeley, 1989. (www.cs.cmu.edu/˜ph)

[32] Stanslav Kovačič and Ruzena Bajcsy, "Chapter 3: Multiscale/Multiresolution Representations," *Brain Warping*, Arthur W. Toga, Ed., Academic Press, San Diego, 1999.

[33] M. Belge, M. Kilmer, and E. L. Miller, "Simultaneous Multiple Regularization Parameter Selection By Means of the L-Hypersurface with Applications to Linear Inverse Problems Posed in the Wavelet Transform Domain," *SPIE International Symposium on Optical Science, Engineering, and Instrumentation: Bayesian Inference for Inverse Problems*, San Diego, July 1998.

[34] Paul Thompson and Arthur W. Toga, “Chapter 18: Anatomically Driven Strategies for High-Dimensional Brain Image Warping and Pathology Detection,” *Brain Warping*, Arthur W. Toga, Ed., Academic Press, San Diego, 1999.

[35] Fred L. Bookstein, “Chapter 10: Linear Methods for Nonlinear Maps,” *Brain Warping*, Arthur W. Toga, Ed., Academic Press, San Diego, 1999.