COMPUTATIONAL ANATOMY: AN EMERGING DISCIPLINE [*]

Ulf Grenander[†] and Michael I. Miller[‡]

June 25, 1998

Abstract

This paper studies mathematical methods in the emerging new discipline of *Computational Anatomy*. Herein we formalize the Brown/Washington University model of anatomy following the global pattern theory introduced in [1, 2], in which anatomies are represented as deformable templates, collections of $0, 1, 2, 3-$dimensional manifolds. Typical structure is carried by the template with the variabilities accommodated via the application of random transformations to the background manifolds. The anatomical model is a quadruple $(\Omega, \mathcal{H}, \mathcal{I}, \mathcal{P})$, the background space $\Omega \doteq \cup_{\alpha} M_{\alpha}$ of $0, 1, 2, 3-$dimensional manifolds, the set of diffeomorphic transformations on the background space $\mathcal{H} : \Omega \leftrightarrow \Omega$, the space of idealized medical imagery $\mathcal{I}$, and $\mathcal{P}$ the family of probability measures on $\mathcal{H}$. The group of diffeomorphic transformations $\mathcal{H}$ is chosen to be rich enough so that a large family of shapes may be generated with the topologies of the template maintained. For *normal anatomy* one deformable template is studied, with $(\Omega, \mathcal{H}, \mathcal{I})$ corresponding to a *homogeneous space* [3], in that it can be completely generated from one of its elements, $\mathcal{I} = \mathcal{H} I_{temp}, I_{temp} \in \mathcal{I}$. For *disease*, a family of templates $\cup_{\alpha} I_{temp}^{\alpha}$ are introduced of perhaps varying dimensional transformation classes. The complete anatomy is is a collection of homogeneous spaces $\mathcal{I}_{total} = \cup_{\alpha}(\mathcal{I}^{\alpha}, \mathcal{H}^{\alpha})$.

There are three principal components to computational anatomy studied herein.

1. *Computation of large deformation maps:* Given any two elements $I, I' \in \mathcal{I}$ in the same homogeneous anatomy $(\Omega, \mathcal{H}, \mathcal{I})$, compute diffeomorphisms h from one anatomy to the other $I \underset{h^{-1}}{\overset{h}{\rightleftarrows}} I'$. This is the principal method by which anatomical structures are understood, transferring the emphasis from the images $I \in \mathcal{I}$ to the structural transformations $h \in \mathcal{H}$ which generate them.

2. *Computation of empirical probability laws:* Given populations of anatomical imagery and diffeomorphisms between them $I \underset{h_n^{-1}}{\overset{h_n}{\rightleftarrows}} I_n, n = 1, \ldots, N$, generate probability laws $P \in \mathcal{P}$ on $\mathcal{H}$ which represent the anatomical variation reflected by the observed popuation of diffeomorphisms $h_n, n = 1, \ldots, N$.

3. *Inference and disease testing:* Within the anatomy $(\Omega, \mathcal{H}, \mathcal{I}, \mathcal{P})$, perform Bayesian classification and testing for disease and anamoly.

KEYWORDS: medical imaging, pattern theory, deformable templates

1   Computational Anatomy: An Emerging Discipline

1.1   Introduction

When the Graduate Division of Applied Mathematics was founded 50 years ago at Brown University, its original research profile was dominated by mechanics: elasticity, plasticity and fluids. Indeed, the founder, Dr. William Prager, was famous as the creator of plasticity theory and the faculty assembled by him had several members highly prominent in the various specialties of mechanics. He later decided that it was necessary for the Division to branch our into other areas of applied mathematics, but mechanics remained its center for a long time.

Therefore it is appropriate for a paper appearing in this Jubilee volume to be related to mechanics. This paper deals with Computational Anatomy, which we view as belonging to the discipline of geometry, governed by pattern theoretic principles, and whose kinematics is described in terms of concepts borrowed from continuum mechanics. In particular we shall show how ideas from elasticity and visco-elasticity can help to represent the biological variability that is the main object of study in Computational Anatomy.

For the past half decade our group has been involved in the development of mathematical and software tools specialized to the representation and inference of gross anatomical structures in the brain [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25]. As the ambition is to create algorithmic tools which help neuroscientists and diagnosticians in the analysis of the substructures of the human brain, the main difficulty we face is that the anatomical substructures form highly complex systems, with variation being the rule! Devising representations and algorithms for inference of brain anatomy in which the subject matter knowledge is precisely and completely specified leads to the development of *anatomically complex knowledge representations* [7].

This area of the mathematical codification of biological and anatomical structure has been exploding over the past several decades. Digital electronic data bases are currently available [26], especially for colocalization of volume datasets such as those encountered with PET/SPECT, CT and MRI [27, 28, 29, 30, 31]. Suitable atlases supporting neuromorphometric analyses [32] are now becoming available with the advent of large volumetric image data sets with large numbers of voxel samples such as CT, MRI and CRYOSECTION [33, 34, 35, 36, 37, 38]. Complementary to the atlas development work there has been great progress on anatomical mapping. Beginning with Bookstein's pioneering work [39], investigators have succeeded in anatomical mapping based on geometric features such as landmarks (points) [39, 40, 41] and contours(lines) [42, 43]. For such approaches predefined subsets of the anatomy provide registration information and become the principal features about which variational studies between the various coordinate systems proceed. The complementary activities on active geometry has presented significant advances in the studies of submanifolds of shape, including the active contour and surface deformation work in the non-statistical setting by the Terzoupoulis school [44, 45, 46, 47, 48], Pentland and Sclaroff [49], and Staib and Duncan [50], as well in the statistical setting the work of Cootes and Taylor [51] and Grenander and Miller [7, 52].

In the whole volume arena, a complementary activity has arisen out of the earliest elegant volume mapping work of Bajcsy and collaborators [28, 29] begun in the early 80's and continuing today [53, 4, 5, 6, 8, 54, 12, 55, 56, 14]. The transformations are applied throughout the continuum of the coordinate system of the template and target, with the voxel image data providing the matching forces throughout the continuum. These approaches have tended to study higher dimensional transformations. Concurrently, the beautiful work by the Evan's group has proceeded along the lines of establishing the power of these approaches in lower and moderate dimensional settings in which the image structures being matched are more globally defined with the variability of global structures being the primary goal of successful registration [36, 37, 57, 58, 59]. Structured approaches studying the variations of substructures via boundaries and surfaces via vector field transformations are now emerging in several groups [9, 60, 61, 62]. As well, detailed gross macroscopic study of cortical folding and localization of functional and anatomical boundaries are emerging in both macaque and human for the Dale and Sereno and Van Essen groups [63, 64, 65, 66, 38, 23]. Via the earliest work of Thirion and Gourdon [67, 68], and more recently Kent and Mardia [69] and Khaneja and Bikircioglu et al. [70, 24], automated methods for generating the curves on brain surfaces are emerging as well.

Because of the shear complexity of human anatomy, in particular the human brain, we have emphasized the study of shape as a collection of submanifolds all taken together forming the complete volume. Therefore, the methodology must combine approaches associated with the active geometry of points, curves, and surfaces, as well as volume mapping. We define our mappings to be groups of diffeomorphisms acting on the background spaces collections of manifolds. Diffeomorphisms on the entire background space when restricted to the submanifolds carry the topologies smoothly as well. Through the recent work of Dupuis et al.[22], a complete understanding has emerged of the general smoothness conditions required for the generation of diffeomorphisms via the transport equation introduced in [13]. This is most akin to the beautiful work emerging from the Azencott school, in particular that of Trouve and Younes. Younes studies diffeomorphic curve matching [71] where a Riemannian metric is defined on a group involving diffeomorphisms which acts transitively on plane curves. Geodesics on the acting group are explicitly computed, and a distance, yielding a variational mathing paradigm, is defined between plane curves by projection (or least action principle). Trouve examines

more general manifolds [72, 73, 74, 75, 76] providing a geometrical framework in terms of infinite dimensional groups. Starting from a metric on the set of smooth vector fields on a compact Riemannian manifold without boundary, Trouve establishes similar conditions as in Dupuis et al. [22] for the existence of the minimizer to the matching problem in the space of diffeomorphisms. It is our belief that this overall approach of studying transformations as groups of diffeomorphisms is the most tenable approach to describing biological shape as complex as human anatomy.

This greater area in which we find ourselves and others working is what we are coming to call the emerging discipline of *computational anatomy*. It is progress in this area which we describe today, emphasizing the work in our own groups at Brown and Washington Universities.

1.2 The Brown/Washington Model of Anatomy

We use the *global shape models* introduced in [1, 2] to represent anatomies as deformable templates. The typical structure is carried by the template; the variabilities are accommodated via the introduction of probabilities on the random transformations. The anatomical model is a quadruple $(\Omega, \mathcal{H}, \mathcal{I}, \mathcal{P})$, the background space $\Omega \doteq \cup_\alpha M_\alpha$ a collection of $0, 1, 2, 3-$dimensional manifolds, the set of diffeomorphic transformations on the background space $\mathcal{H} : \Omega \leftrightarrow \Omega$, the space of ideal imagery $\mathcal{I}$, and $\mathcal{P}$ the family of probability measures on $\mathcal{H}$.

The group of diffeomorphic transformations $\mathcal{H}$ are constructed to be rich enough so that a large family of shapes may be generated with the global topologies of the template maintained. We begin with the basic assumption that there is essentially only one deformable template, i.e. one anatomical ensemble $(\Omega, \mathcal{H}, \mathcal{I}, \mathcal{P})$ which forms a a *homogeneous space* [3], or what we shall term a *homogeneous anatomy*. The basic property is that it can be completely generated from one of its elements, $\mathcal{I} = \mathcal{H} I_{temp}, I_{temp} \in \mathcal{I}$. In the homogeneous anatomy setting, all elements $I \in \mathcal{I}$ are essentially the same modulo a diffeomorphism in $\mathcal{H}$. More generally, as for the disease state, an area that has received less attention [77] then the study of normal anatomy,, we shall require a family of templates $\cup_\alpha I_{temp}^\alpha$ with perhaps varying dimensional transformation classes. The complete anatomy is built from collections of homogeneous spaces $\mathcal{I}_{total} = \cup_\alpha (\mathcal{I}^\alpha, \mathcal{H}^\alpha)$. We return to this below.

Within this framework, we see that there are three principal components to computational anatomy:

1. **Computation of large deformation maps:** Given any two elements $I_n, I_m \in \mathcal{I}$ in the same homogeneous anatomy $(\Omega, \mathcal{H}, \mathcal{I}, \mathcal{P})$, compute diffeomorphisms h with inverses $h^{-1} \doteq \phi$ from one anatomy to the other $I_n \underset{\phi_{nm} \doteq h_{nm}^{-1}}{\overset{h_{nm}}{\rightleftarrows}} I_m$. This is the principal method by which anatomical structures are understood, transferring the emphasis from the images $I \in \mathcal{I}$ to the structural transformations which generate them.

2. **Computation of empirical probability laws:** Assume we are given populations of anatomical imagery and diffeomorphisms between them $I_n \underset{\phi_{nm} \doteq h_{nm}^{-1}}{\overset{h_{nm}}{\rightleftarrows}} I_m, n, m = 1, \ldots, N \subset \mathcal{H}$. The second component of computational anatomy calculates probability laws $P \in \mathcal{P}$ on $\mathcal{H}$ which represent anatomical variation reflected by the observed popuation.

3. **Computational inference on population and disease testing and classification:** Given a anatomical representation in the form of prior probability measures on the deformable template representing the anatomy, perform Bayesian classification and testing for disease and anamoly.

1.3 Generating Large Deformation Maps

The large deformation maps generated in our group are constructed using methods which have been under development over the past 6 years [4, 5, 6, 9, 11, 13, 5, 16, 17, 20, 18, 22]. These approaches generate vector fields which are bijective maps between families of templates and targets, and are based on large deformation kinematic models. The maps are of high dimension, allowing for the local dilation and contraction of the coordinates of the template into the coordinates of the individual anatomy accommodating the very finest details in variation. The large deformation diffeomorphisms are constructed as the solution of ordinary differential equation corresponding to the transport equation in mechanics. The deformations thus correspond to flows; the fact that they are integrated in time guarantee their large deformation diffeomorphic character. Our purpose in constructing such *invertible* maps is that they can be used to study geometric features of the target anatomies. The transformation are constrained to be 1-1 and onto, and differentiable with differentiable inverse, so that connected sets remain connected in the target, surfaces are mapped as surfaces, and the global relationships between structures are maintained. For example, connected such as *ventricles, thalami, deep nuclei* in brain anatomy all remain connected under the mapping. As well, the large deformation maps carry the coordinates of disparate families onto each other while maintaining the geometric features. Thus they do not bias the empirical statistic building towards one brain or the other. This would be in sharp contrast to, for example, large

deformation elasticity as used in [54], in which the fundamental assumption associates energy to physical distance and size of deformation between objects. For mapping a single brain to itself during growth, development and aging, this would be more appropriate; mapping across brains seems to require a deformation setting in which one is not penalized in the same manner.

In our work, not only are the global structures of *deep nuclei* important but as well we study the differential geometric features associated with the finest geometric structures including *sulcal trajectories and cortical folds*. Over the past several years we have been involved in a collaboration with David Van Essen as he studies the cortex of the macaque and the Visible Human in cryosection [78, 10, 65, 23]. The awesome precision revealed by such data sets supports natural biological questions concerning geometric features. Notions such as Riemannian length, Gaussian curvature, and surface area measures of highly complex folded structures are at the heart of our investigative work and others [79]. Methods which allow for the study of shape in a quantitatively meaningful manner are becoming essential. It is therefore natural to organize the transformations around the continuum, emphasizing the properties of diffeomorphisms as they map the various tangent spaces and curvature features of the embedded submanifolds. Instead of pixel based approaches we prefer to base the analysis on the Euclidean space in which the brain structures actually reside. In this approach, the maps are used to analytically compute classical formulas from differential geometry on transformation of the tangent spaces and curvature tensor under the Jacobian and Hessian. For this the transformations must be established on the continuum. We deal with the awesome dimension of the transformation by organizing them into a hierarchy of subgroups, with dimension which increases with successive subgroups [17]. At the highest dimension $(>> 10^7 - 10^8)$, the transformations are rich enough to map the finest anatomical details. If fiducial markers, landmarks and surface information are known then they can be consistently combined with the partial differential equation (PDE) representations of the variational solutions representing the optimal transformations (see [9], for example). Ultimately, the maps are based on the full continuum of coordinates with embedded points, lines, surfaces, and subvolumes carried into the target by the high dimensional smooth volume transformations.

1.4 Empirical probability construction and complex priors for image reconstruction

The Bayesian approach to understanding complex brain variability is to construct one or several templates containing the topological structures of complex anatomy. Upon this a probability measure is constructed on the space of $I\!R^3$ valued vector field transformations $\mathcal{H}$. This is the knowledge representation of biological variability. The prior measure encodes probabilistic properties of the maps which are required for understanding the various differential geometric invariants and physical features such as distances, surfaces, and volumes in the infinite family of possible anatomies. The template and the prior measure precisely specify the global anatomical relationships between structures as well as how they can vary from one brain to another.

For constructing such knowledge representations empirically, the template and the parameters of the prior measure are directly estimated from subpopulations of brain anatomies. This is similar to the approach taken in [7] for the representation of *mitochondria/membrane* organelle shape. Organelle templates and prior measures were constructed by mapping 100's of biological shapes to topologically equivalent structures. Such an empirically based procedure is proposed herein, requiring the careful definition of the template via the family of maps to the population of anatomical shapes which the template represents. The empirical average of the family of maps provides the minimum mean-squared error (M.M.S.E.) template. It is important to realize that the average should be taken in transformation space, not in image space. The empirical variation of the family of transformation fields provides the information for the probabilistic properties of the prior measure defining the space of vector field transformations. Prior estimation corresponds to *generalized spectrum estimation*.

One of the principal motivations of our work is associated with several longstanding involvements in medical imaging on the development of *anatomically complex prior distributions* for probabilistic image reconstruction. For imaging modalities arising in positron emission and single photon emission tomography (PET/SPECT), in contrast to computed tomography (CT) and magnetic resonance imaging (MRI), probabilistic image reconstruction algorithms are required due to the inherently low signal-to-noise ratios (finite photon counts). Since the earliest introduction of the EM algorithm by Shepp and Vardi in 1982, the 80's and 90's have seen the development of an extensive literature in maximum likelihood (ML) reconstruction [80, 81, 82, 83, 84, 85, 86, 87, 88, 89] incorporating the fundamentally statistical nature of the measurement process. In the early 80's, introduction into the ML formulation of prior distributions on the image class was resisted. With the identification of the PET/SPECT reconstruction problem as a density estimation problem in function space, and the subsequent connection to Grenander's demonstration on the inconsistency of maximum-likelihood from finite data sets [90], there has been a rapid influx of prior based regularized MAP reconstruction methods [91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101], including the method of sieves, Good's roughness, Gaussian and Gibb's random fields, as well as prior distributions controlled by edge and boundary processes. In all of these approaches the prior is anatomically simple and relatively uninformative. It reflects little if any *a-priori* information about the individual anatomy itself.

Modern neuroimaging methods have now reached the point where exquisitely detailed *in vivo* information regarding the anatomical structure of each individual brain is routinely collected. Commonly individuals receive multiple modality scans including CT and MRI providing relatively precise $< (1mm)^3$ anatomical resolution. Such anatomical 'side information' provides significant constraints and enhancement for functional imaging studies. Studies of cerebral hemodynamic fluctuation, derived during various mental task activation studies [102, 103, 104, 105], for example, generally provide substantlly lower anatomical resolution. There has as well been a growing body of work on the introduction of more structured models for incorporation of side information [106, 107, 58]. To date the incorporation of anatomical information has been hindered by the inability to expeditiously relate such information between morphologically varying brains. The difficulty lies in two areas. First images between differing anatomies must be registered. Second, even when registered, normal variation across disparate anatomies makes pooling of inter anatomical data difficult, if not impossible. Providing mathematical tools which allow for structural decompositions of individualized brain tissue would substantially increase the effectiveness of reconstructions which are constrained via the individual's own anatomy (see Figure 7 of [4] and [108], for example).

1.5   Normal versus disease state.

Our goal is to achieve *understanding* of anatomical variability in the sense that we can make precise probabilistic statements about the variability: this is what we mean by *image understanding*. This plunges us deeply into the study of normal versus disease as indicated via anatomical variation. Almost all of our discussion has focussed on *normal* anatomical structure. Concurrently, we and co-workers are developing anatomical atlases with the intention of providing a systematic approach for the study of the neuropsychiatric disease [109] and associated shape change [55, 56] as well as for use in plastic surgery [110, 111]. To date in the study of morphometric change and its link to neuropsychiatric disease, manual and semi-automated methods for volumetric assessment of small substructures within the brain are relatively error prone with the reliability of these methods only barely adequate to make quantitative statements about disease (see for example applications in schizophrenia [112, 113, 114]). More importantly, most current segmentation based methods for morphometry do not capture information about the shape of structures well. While the volumes of brain structures and substructures may be related to aspects of the pathogenesis of disease, critical aspects of disease may not be reflected in simple volume measures alone. There is already evidence that the qualitative feature of morphometric shape, such as the patterning of sulci and gyri on the cortical surface [115], are indicators of disease. The precise and quantitative assessment of shape characteristics of volumes and surfaces could conceivably offer insights into a multitude of critical questions.

A serious discussion of variation away from normal plunges us immediately into the very heart of Bayesian hypothesis testing. For years people such as Marcus Raichle in cognitive science and neurology have argued for the necessity of a more complete understanding of the normal thinking brain as tantamount to the understanding of disease state. This is fundamental to the Bayesian paradigm. The base measure representing normal anatomical variation (the null hypothesis) is the basic reference measure. Populations of patients with neuropsychiatric diseases, characterized by subtle and distributed abnormalities of brain structure and shape can only be compared to normal controls through the prior measure reflecting the statistical summary of normal and abnormal variation of the component structures of the brain. Towards this end, we introduce the space of normal anatomies $(\mathcal{I}, \mathcal{H}, P)$, $\mathcal{I}$ the image algebra of all possible anatomical images associated with transformations on the coordinate system $\mathcal{H} : \Omega \leftrightarrow \Omega$. The image algebra will be a union $\mathcal{I}_{total} = \cup_\alpha (\mathcal{I}^\alpha, \mathcal{H})$ where all images in $\mathcal{I}^\alpha$ are topologically equivalent. Characterizing normals becomes an empirical procedure of constructing probability laws P from the family of transformation $\mathcal{H}$ observed in actual anatomies. Thus far in our work, *disease* or *abnormality* refers to either of two kinds of anatomical variation: (i) local and regional alterations in the size and shape of brain structures relative to the normal base measure, i.e. a *large deviation from normal*, or (ii) the abnormal presence (e.g., neoplasm) or absence (e.g., congenital) of substructures corresponding to a fundamental change in topology of the underlying graph structure of the brain.

In the first case (i), the disease state corresponds to a transformation $h \in \mathcal{H}$ of one anatomy to another. The signalling of disease, therefore, corresponds to a transformation which is a *large deviation* from the identity in the group of transformations as reflected by the normal and disease measures (P_0, P_1), with the disease measure absolutely continuous with respect to the normal $P_1 >> P_0$. For example, for characterizing disease the image template representing normal $I_{temp} \in \mathcal{I}$ can have a hippocampus that is classified as normal, while there could exist a transformation $h \in \mathcal{H}$ at considerable distance from the unit element of the transformation group. The subcomponents of the image could be so small that it may indicate disease such as ventricular enlargement, hippocampul and or cortical shrinkage such as might be associated with various neuropsychiatric diseases, alzheimer's, alcoholism, or schizophrenia. To quantify such statements we use the prior P on $\mathcal{H}$.

This brings us to the second aspect of disease, (ii) in which the topology is fundamentally different. This requires the extension to transformations that destroy topologies, *diffeomorphisms with tears* to use David Mumford's terminology. This is an ambitious area of current research. The relation between homogeneous and normal anatomies is not completely

straightforward. On the one hand, all the images in a homogeneous anatomy generated by a normal template $\mathcal{I}_{temp}^{\alpha}$ need not be normal. In contrast, the set of normal anatomies need not constitute a homogeneous anatomy, since one can well imagine two images classified as normal although they are not topologically equivalent. This leads to the dichotomy of transformations as well, $(\mathcal{H}^{a}, \mathcal{H}^{h})$ corresponding to *automorphic* or topology preserving deformations versus *heteromorphic*, topology destroying deformations. Most of the effort within this paper is focussed on topology preserving transformations. In the final sections topology destroying transformations are presented that annihilate or create certain anatomical substructures.

1.6 Paper layout.

In section 2 the anatomical model is presented as an orbit under the group action of diffeomorphisms with a distance metric placed on the orbit. With this distance metric, the large deformation landmark and image matching problems are presented. Section 3 examines examples of large deformation brain mapping, studying the neocortex and hippocampul subvolumes. In section 4 small deformations are examined, with small deformation templates for entire brains in the macaque, and hippocampus surfaces in the human. Sections 5 and 6 examine small and large deformation probability measures for anatomical representation as generalized autoregressive modelling on Gaussian fields. Sections 7 and 8 examine large deformation surface and curve matching. The distances between surfaces is associated with the Hilbert Schmidt distance between the shape operator curvature representation of the surface; for curves the Hilbert Schmidt distance between the orthogonal Frenet curvature and torsion representation of curves is used. Bipartite graph matching is presented for the curve matching problem, with prior distributions induced via the Frenet representation of curves examined as well. Section 9 examines the study of symmetry through the orthogonal group, and section 10 examines disease introducing two models of disease, automorphic and heteromorphic patholigies, the former topology preserving and the latter topology changing.

2 The Anatomic Model

Restrict attention to brain anatomy. The template brain coordinate system $\Omega \doteq \cup_{\alpha} M_{\alpha} = [0,1]^3$ is a collection of heteregeneous submanifolds all of which have been selected as *biologically meaningful since they are biologically stable*. Figure 1 shows various examples depicted via volume tissue sections corresponding to the whole brains from the macaque monkey from David Van Essen's laboratory. Top left panel shows the surface rendering of $640 \times 480 \times 200$ voxel volume, with the superimposed sulcal lines (contours). The top right panel shows a section through the entire volume. Notice the beautiful cortical folds, the deep nuclei, as well as the hippocampus. The bottom left panel shows a rendering of the C^2 manifold representing the cortical surface consisting of locally quadratic charts of the full triangulated graph of the surface (see Joshi et al. [10]). The bottom right panel shows a section through a brain cryosection with the cortical surface embedded in the volume delineating the interface between the gray and the white matter.

2.1 The Anatomical Orbit

Variability is studied through the transformations $h \in \mathcal{H} : \Omega \subset I\!\!R^3 \leftrightarrow \Omega$ the space of homeomorphisms constructed from vector fields

$$h : x \doteq (x_1, x_2, x_3) \in \Omega \mapsto h(x) \doteq \underbrace{(h_1(x), h_2(x), h_3(x))}_{\doteq x - u(x)} \in \Omega \ , \tag{1}$$

where it is sometimes convenient to study the transformations modulo the identity map in terms of the $u(\cdot)$ field. The negative sign gives the interpretation as the Eulerian displacement of the particulate description of flow; a particle at position x originated at point $x - u(x)$ in the original coordinates. Anatomic variation is studied through the inverse maps as well, $\phi \doteq h^{-1} \in \mathcal{H}$. As we choose to calculate arc length, surface areas, and volume measures we require the transformations to be diffeomorphic. The transformations are assumed rich enough to carry a single element through the space of anatomies. Now consider the ensemble of anatomical imagery $\mathcal{I}$ the image algebra of functions $I \in \mathcal{I} : \Omega \to \mathcal{R}^M$ defined on Ω the background space on which the coordinate systems are defined:

$$I \doteq \{I(x), x \in \Omega\} \ , \ \Omega \doteq [0,1]^3 \doteq \cup_{\alpha} M_{\alpha} \subset I\!\!R^3 \ .$$

The range or contrast space $\mathcal{R}$ in which the measured imagery takes its I-values are in general M-vectors with components corresponding to intensities from various sensor modalities, cell types, names of brain nuclei, etc. Throughout they shall correspond to the idealized way in which imagery of MRI and the gray level luminance of gross cryo-sections (CRYO) as measured via CCD detection would appear without noise. Thus, the imagery corresponds to the scalar mean fields
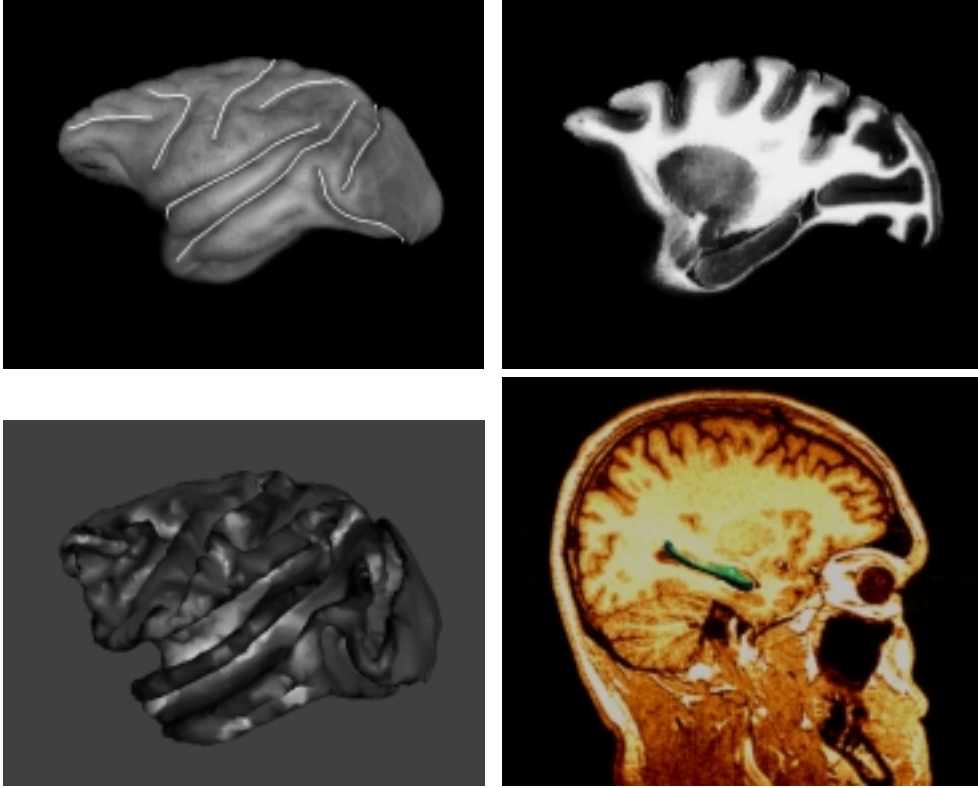
Figure 1: Top left panel shows a macaque brain hemisphere with some of the sulcus fundus curve paths depicted. The top right panels shows a whole brain section. The bottom left panel shows the C^2 manifold representing the cortical surface at the interface of gray and white matter. The bottom right panel shows a section through a whole brain MRI with the hippocampul surface embedded. Data were taken from a whole brain volume reconstructions from the laboratory of David Van Essen of the Department of Anatomy and Neurobiology and John Csernansky of the Department of Psychiatry both of Washington University.

which parameterize the measurements in the idealized case, with no noise or blur in the optical transfer. The group of diffeomorphisms divides the anatomical ensemble into discjoint orbits.

Theorem 1 *Given the group of transformations $(\mathcal{H}, \circ)$ with law of composition $\circ$ and identity element e, then $\mathcal{H}$ acting on $\mathcal{I}$ defines a group action $G : \mathcal{H} \times \mathcal{I} \to \mathcal{I}$ according to*

$$G(h, I) \quad \doteq \quad hI \doteq \{I(h(x)), x \in \Omega\} \; , \quad \text{where } G(h_2, G(h_1, I)) = G(h_2 \circ h_1, I) \doteq h_2 \circ h_1 I \; . \tag{2}$$

Then $\mathcal{I}$ is composed of a set of disjoint orbits which are each homogeneous spaces under the group action of diffeomorphisms:

$$\mathcal{I} = \cup_\alpha \mathcal{H}^\alpha I_{temp}^\alpha \tag{3}$$

Proof: G as defined is a group action, and therefore defines an equivalence relation $I_1 \sim I_2$ if $\exists h \in \mathcal{H}$ such that $I_1 = hI_2$ (see [3]). Since it is an equivalence relation, it divides the space of anatomies into disjoint orbits. ∎

It is assumed throughout most of the paper that *the anatomical ensemble $\mathcal{I}$ is a homogeneous space*, so that there is only one template I_{temp}, and the set of diffeomorphic transformations $\mathcal{H} : \Omega \leftrightarrow \Omega$ associated with the image algebra $\mathcal{I}$ act transitively on $\mathcal{I}$. That is, any two images $I_1, I_2 \in \mathcal{I}$ are topologically equivalent in the sense that

$$\exists h \in \mathcal{H} \ni I_1 = hI_2 \text{ where } hI \doteq \{I(h(x)), x \in \Omega\} \; . \tag{4}$$

This makes precise the notion of a deformable template, one selected and fixed anatomy $I_{temp} \in \mathcal{I}$ such that $\mathcal{I} = \mathcal{H}I_{temp} = \{hI_{temp} : h \in \mathcal{H}\}$, and $(\mathcal{I}, \mathcal{H})$ is the orbit under $\mathcal{H}$ of the template I_{temp}. The template $I_{temp} \doteq \{I_{temp}(x), x \in \Omega\}$ and the associated imagery $\{I^{(1)}, I^{(2)}, \dots I^{(M)}\}$ represent the mean fields associated with each of the M-imaging modalities.

For purposes of understanding the populations of anatomy, modulo the special Euclidean group $\mathbf{E}(3) = \mathbb{R}^3 \otimes \mathbf{SO}(3)$ (semi-direct product of the translation and generalized linear group), the template itself carries the affine motions with it. The template is then a pattern or orbit of I_{temp} under the rigid motion. Throughout this distinction is assumed understood, with all of the diffeomorphic maps $\mathcal{H}$ generated modulo this rigid motion. In our setup a pattern is an equivalence class of images from $\mathcal{I}$ with respect to a similarity group $\mathcal{S}$, for example $\mathcal{S} = \mathbf{E}(3)$ (see [1, 2]).

2.2 Distances between Images in the Orbit

The brain data shown in Figure 1 provides specific motivation for the fact that the complexity of the patterns which must be represented—ventricles, cortical folds, etc.—requires transformations which are of high dimension. However, all quadratically based penalties such as used in [116, 29, 117, 4] have restoring forces to the template proportional to the distance displaced. Large magnitude displacements are severely penalized. To accommodate large magnitude deformations, small deformation assumptions upon which linear models and quadratic penalties are based (see [118, 119]) are relaxed, motivating the formulation of transformations in an Eulerian setting incorporating the *large deformation kinematics* of Christensen et al. [13, 120, 6, 5]. In this setting the restoring force which normally grows with deformation length is allowed to relax with *time*, originally Rabbitt's formulation [5], while still maintaining smoothness of the deformation field. This physical principal becomes the basis for our proposed metric distance between shapes.

To map one anatomy to another we need to define a distance, or a metric $\rho(I, I'), I, I' \in \mathcal{I}$ between elements assumed from the same orbit. We do this assuming that the diffeomorphisms between the image pair $I \underset{\phi(T) \doteq h^{-1}(T)}{\overset{h(T)}{\rightleftarrows}} I'$, evolve in time as a flow given by the solution to the transport ordinary differential equation (ODE). Defining the operator $\nabla_x^t \doteq [\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}\frac{\partial}{\partial x_3}]$, then the diffeomorphisms are identified with the end point condition of $t \in [0, T]$ given by

$$
\begin{aligned}
\frac{d\phi(x,t)}{dt} &= v(\phi(x,t),t) \ , \ \phi(x,0) = x \ , \\
\frac{\partial h(x,t)}{\partial t} &= -\nabla_x^t h(x,t) v(x,t) \ , \ h(x,0) = x \ .
\end{aligned}
\tag{5}
$$

We identify the diffemorphisms with the space time velocity field $v(\cdot, \cdot) : \Omega \times [0, T] \rightarrow \mathbb{R}^3$ corresponding to the tangent field of the flow. The distance beween two images is determined by the distance from the origin (identity map) of the associated velocity field which generates it. Following Dupuis et al. [18, 22] the space of velocity fields is assumed to be in a Hilbert space $\mathcal{V}$ with norm $\| \cdot \|_\mathcal{V}$ and inner product $< \cdot, \cdot >_\mathcal{V}$. To ensure that the ODE Eqn. 5 has a well defined solution it is constrained to be an element $v \in \mathcal{V}$ corresponding to a Sobelev space whose inner product and norm corresponds to a fixed number of derivatives. The order of differentiation ensures sufficient smoothness so that the solution of the ODE is unique and defines a diffeomorphism.

In all of the matching work [5, 6, 11, 13], we have chosen the order of differentiablity to correspond to differential operators from continuum and fluid mechanics, so that a positive definite linear differential operator of the form $L = -a\nabla^2 + b\nabla\nabla \cdot +c$ is chosen, which induces the norm in $\mathcal{V}$ according to $\|v\|_\mathcal{V}^2 = \int_{[0,T]} \int_\Omega \|Lv(x,t)\|^2 dx dt < \infty$. We have used operators which differentiate only in space. The form of the operator is made explicit as we go.

Definition 1 *Define the distance* $\rho(\cdot, \cdot) : \mathcal{I} \times \mathcal{I} \rightarrow \mathbb{R}^+$ *between elements* $I, I' \in \mathcal{I}$ *to be*

$$
\begin{aligned}
\rho(I, I') &\doteq \min_{v \in \mathcal{V}} \int_0^T \int_\Omega \|Lv(x,t)\|^2 dx dt \text{ subject to } h(0)I = I, \ h(T)I = I' \\
&\qquad \text{where } h(x,T) = \int_0^T -\nabla_x^t h(x,t) v(x,t) dt + x \ , x \in \Omega \ ; \\
&\doteq \infty \qquad \text{if } \nexists v \in \mathcal{V} : I' = h(T)I \ .
\end{aligned}
\tag{6}
$$
$$
\tag{7}
$$

It is sometimes convenient to denote the final time diffeomorphism as simply $h(\cdot) \doteq h(\cdot, T)$. That this properly defines a distance follows from the facts that (1) it is symmetric, and (2) satisfies the triangle inequality. The symmetric property follows since $v \in \mathcal{V}$ being sufficiently smooth implies the ODE has a unique solution, with $\phi = h^{-1}$ given by the solution driven by the negative of the velocity field, flowing opposite in time. Thus $\rho(I, I') = \rho(I', I)$. The triangle property follows from the Hilbert space distance being determined by the norm-squared $\|Lv\|^2$.

It would be nice to know that all distances between images are not infinity, i.e. take elements I, I' which are diffeomorphic, then there exists velocity fields of finite energy which when integrated through the ODE generate the diffeomorphism. At least for the finite dimensional landmark matching problem, this is the case. As proven in Joshi via the homogeneity lemma [16], given anatomies I, I' contained in a compact subset of $(\mathbb{R}^3)^L$, then there exists finite energy velocity fields which generate diffeomorphisms from one to another.

2.3 Landmark and Image Matching

Brain geometry is studied through the diffeomorphic maps from one anatomy to another. For generating large deformation diffeomorphisms we rely on the basic protocol described in [5, 6, 9, 11, 13, 16, 17, 18, 19, 22] in which maps of increasingly higher dimensions are generated one after another through composition [17]. This combines information at finer scale by composing the large deformation diffeomorphic transformations $h = h_n \circ \cdots h_2 \circ h_1$. Successive combinations of transformations are chosen including the affine motions, rigid motions generated from subgroups of the generalized linear group, large deformation landmark transformations and the high dimensional large deformation image matching transformations. The dimension of the vector field transformations are successively being increased. Given the distance between imagery in the orbit $\mathcal{I}$ the matching problem between elements I, I' is straightforwardly defined. Since all of the imagery being matched are observed with noise, model them as conditional Gaussian random fields. Take I as the provisory template; then the measured imagery I' is a conditionally Gaussian random field with mean field the provisory template composed with the unknown diffeomorphism hI, and fixed variances. The diffeomorphism $h, \phi \doteq h^{-1}$ are identified with a velocity field $v \in \mathcal{V}$ in the Hilbert space of space-time fields. The problem is to estimate the velocity field which matches I to the observable image I' with finite penalty.

Assume each *target* is characterized via a set of landmarked imagery $I \doteq \{x_l \in \Omega, l = 1, 2, \cdots, L\}$ and $I' \doteq \{y_l \in \Omega, l = 1, 2, \cdots, L\}$ or digitized cryosection imagery (CRYO) or magnetic resonance imagery (MRI) $I \doteq \{I(x), x \in \Omega\}$, $I' \doteq \{I'(x), x \in \Omega\}$. Throughout a quadratic distance function $\|hI - I'\|^2$ is used defining the matching problem which attains its minima when the landmarks and or imagery are mapped from the targets back to the template. As well we assume that the anatomical imagery is smooth so that variation such as gradients can be computed and are smooth themselves.

Problem Statement 1 (Large Deformation Matching) *The optimal match of I to observation I' is given by*

$$\hat{h}(x, T) = \int_0^T -\nabla_x^t \hat{h}(x, t)\hat{v}(\cdot, t)dt + x \text{ where } \hat{v}(\cdot) = \arg\min_{v \in \mathcal{V}} \|v\|_{\mathcal{V}}^2 + \|h(T)I - I'\|^2 . \tag{8}$$

The costs for landmark and image matching become

$$landmark \; matching \quad \|hI - I'\|^2 \doteq \underbrace{\sum_{l=1}^{L}[y_l - h(x_l, T)]^t \Sigma_l^{-1}[y_l - h(x_l, T)]}_{D_1(h)} , \tag{9}$$

$$image \; matching \quad \|hI - I'\|^2 \doteq \underbrace{\frac{1}{\sigma^2}\int_\Omega |I(h(x, T)) - I'(x)|^2 dx}_{D_2(h)} . \tag{10}$$

The $\Sigma_l, l = 1, \ldots, L$ are positive definite 3×3 covariance matrices expressing uncertainty in landmark location; σ^2 expresses the noise properties for the image matching problem. The limiting case for equality matching results from sending covariance to zero as demonstrated by Joshi [16]. This is included below.

3 Algorithms for Large Deformation Diffeomorphisms

3.1 Gradient Algorithm for Image Matching

The energetics $\int_0^T \int_\Omega \|Lv(x, t)\|^2 dx dt$ on the space of diffeomorphisms are induced with $L = (-a\triangle - b\nabla\nabla \cdot + cI)$ and powers L^m, $m \geq 1$. Christensen [120] has examined various boundary conditions, zero, mixed and periodic. Since we identify the anatomy with a point $v \in \mathcal{V}$ in the Hilbert space, it is convenient to work with the complete orthonormal representation via a basis $\{\psi_j\}$ of $\mathcal{V}$ expanding the velocity fields in linear combinations

$$v(x, t) = \sum_j \nu_j \psi_j(x, t) . \tag{11}$$

Since the velocity fields are real, the ν, ψ can be constructed to be real which is assumed throughout. Since linear expansions of $v \in \mathcal{V}$ identify the velocity fields with the expansion coefficients of the basis $\{\psi_j\}$, this forms the basis for the gradient computations associated with deterministic Newton descent or the drift terms analogous to the small deformation optimizations described in [117, 4].

Algorithm 1 (Gradient Algorithm) *Define the variation in terms of the basis representation*

$$\hat{h}(x,T) \;=\; \int_0^T -\nabla_x^t \hat{h}(x,t)\hat{v}(x,t)dt + x \text{ where } \hat{v}(x,t) = \sum_{j=0}^{\infty} \hat{\nu}_j \psi_j(x,t) \;, \tag{12}$$

$$\text{and } \hat{\nu}_j, j = 0,1,\ldots \;=\; \arg \min_{\nu:v=\sum_j \nu_j \psi_j \in \mathcal{V}} \int_0^T \int_\Omega \|Lv(x,t)\|^2 dxdt + \|h(T)I - I'\|^2 \;. \tag{13}$$

The gradient algorithm becomes initialize with $n = 0$, $\nu^{(0)} = 0$, *and iterate* $n = 0,1,\ldots$:

1. Calculate transformation:

$$h^{(n)}(x,T) \;=\; \int_0^T -\nabla_x^t h^{(n)}(x,t)v^{(n)}(x,t)dt + x \text{ where } v^{(n)}(x,t) = \sum_{j=0}^{\infty} \nu_j^{(n)} \psi_j(x,t) \;. \tag{14}$$

2. Calculate gradient perturbation:

$$\nu_j^{(n+1)} \;=\; \nu_j^{(n)} - \Delta \left(\int_\Omega \int_0^T L^\dagger Lv^{(n)}(x,t)\psi_j(x,t)dxdt + \frac{\partial}{\partial \nu_j}\|hI - I'\|^2 \bigg|_{\substack{h=h^{(n)} \\ v=v^{(n)}}} \right) \tag{15}$$

where for landmark and image matching

$$\textit{landmark matching} \quad \frac{\partial}{\partial \nu_j}\|hI - I'\|^2 \;\doteq\; \underbrace{2\sum_{l=1}^{L}[y_l - h(x_l,T)]^t \Sigma_l^{-1} \frac{\partial h(x_l,T)}{\partial \nu_j}}_{\frac{\partial}{\partial \nu_j}D_1(h)} \;,$$

$$\textit{image matching} \quad \frac{\partial}{\partial \nu_j}\|hI - I'\|^2 \;\doteq\; \underbrace{\frac{2}{\sigma^2}\int_\Omega (I(h(x,T)) - I'(x))\nabla I(h(x,T)) \cdot \frac{\partial h(x,T)}{\partial \nu_j}dx}_{\frac{\partial}{\partial \nu_j}D_2(h)} \;,$$

with

$$\frac{\partial h(x,T)}{\partial \nu_j} \;=\; \int_0^T -\nabla_x^t h(x,t)\psi_j(x,t)dt + \int_0^T -\nabla_x^t \frac{\partial h(x,t)}{\partial \nu_j}v(x,t)dt \;.$$

After stopping, define the final iterate as $\hat{v} \doteq v^{(n+1)}$ *given by* $\hat{v}(x,t) \doteq \sum_j \hat{\nu}_j^{(n+1)}\psi_j(x,t)$, *with* $\hat{h}(x,T)$ *defined by Eqn. 12.*

Stable points satisfy the necessary maximizer conditions for the optimal transformation optimizing Eqns. 12, 13.

The endpoint variational condition on the diffeomorphism $\frac{\partial h(T)}{\partial \nu_j}$ follows from the standard perturbation argument assuming sufficient smoothness. Define the perturbations

$$h(x,T) \;=\; \int_0^T -\nabla_x^t h(x,t)v(x,t)dt + x \text{ where } v(x,t) = \sum_{j'} \nu_{j'}\psi_{j'}(x,t) \tag{16}$$

$$h_\epsilon(x,T) \;=\; \int_0^T -\nabla_x^t h_\epsilon(x,t)v_\epsilon(x,t)dt + x \text{ where } v_\epsilon(x,t) = \sum_{j'} \nu_{j'}\psi_{j'}(x,t) + \epsilon\psi_j(x,t) \;. \tag{17}$$

The partial derivative at the end point condition becomes

$$\frac{\partial h(x,T)}{\partial \nu_j} \doteq \lim_{\epsilon \to 0} \frac{1}{\epsilon}\frac{h_\epsilon(x,T) - h(x,T)}{\epsilon} \;=\; \lim_{\epsilon \to 0}\frac{1}{\epsilon}\left(\int_0^T -\nabla_x^t h_\epsilon(x,t)v_\epsilon(x,t)dt - \int_0^T -\nabla_x^t h(x,t)v(x,t)dt \right) \tag{18}$$

$$\;=\; \int_0^T -\nabla_x^t h(x,t)\psi_j(x,t)dt + \int_0^T -\nabla_x^t \frac{\partial h(x,t)}{\partial \nu_j}v(x,t)dt \;. \tag{19}$$

Notice the fundamental role of the transport ODE in the variation of the distance term.

Remark 1 Lattice and eigenfunction implementation. *Wang solves the PDE numerically on the lattice using discrete approximations to the differential operators in space. Divide $\Omega \times [0,T]$ into a discrete lattice $(\Omega \times [0,T])^\Delta$ of pixels of fixed size, then $\psi_{jk}(x,t)$ has it's support on lattice site $(x_j, t_k) \in (\Omega \times [0,T])^\Delta$. Define L^Δ to be the discretized version of $L = -a\nabla^2 - b\nabla \cdot \nabla + cI$ with difference approximations for the Laplacian, gradient, and divergence. Wang defines the operator to have no derivatives in time. The locality of the basis reduces the inner product to operations on the lattice sites*

$$\int_{\Omega \times [0,T]} L^\dagger L v(x,t) \psi_{j,k}(x,t) dx dt = L^{\Delta^\dagger} L^\Delta v(x_j, t_k) \ .$$

Since we have defined the norm through the differential operator L, $\|v\|_\mathcal{V}^2 = \|Lv\|^2$, choose ψ, λ to be the eigenelements of L, $L\psi = \lambda\psi$. In general the eigenfunctions are complex. For expansions of real-valued fields, the expansions reduce to a real expansion in sines and cosines determined by the various boundary conditions [4, 120]. Then linear expansions of $v \in \mathcal{V}$ identify the velocity fields with the expansion coefficients $v(x,t) = \sum_{j=0}^\infty \nu_j \psi_j(x,t)$. This makes the first cost term particularly simple,

$$\int_\Omega \|Lv(x,t)\|^2 dx = \sum_j |\lambda_j|^2 |\nu_j|^2 \ .$$

The variations of the cost in the gradient algorithm Eqn. 15 becomes as follows for the lattice and eigenfunction implementsions:

$$lattice \quad \nu_{jk}^{(n+1)} \quad = \quad \nu_{jk}^{(n)} - \Delta \left(L^{\Delta^\dagger} L^\Delta v^{(n)}(x_j, t_k) + \frac{\partial}{\partial \nu_{jk}} \|hI - I'\|^2 \Big|_{\substack{h=h^{(n)} \\ v=v^{(n)}}} \right) \ . \tag{20}$$

$$eigenfunction \quad \nu_j^{(n+1)} \quad = \quad \nu_j^{(n)} - \Delta \left(2|\lambda_j|^2 \nu_j^{(n)} + \frac{\partial}{\partial \nu_j} \|hI - I'\|^2 \Big|_{\substack{h=h^{(n)} \\ v=v^{(n)}}} \right) \ . \tag{21}$$

3.2   Greedy Algorithm for Image Matching

Christensen [4, 5, 6, 11, 13] has implemented a computationally efficient algorithm for image matching which exploits the fact that the operator does not differentiate in time. Discretize space-time $\Omega \times T$ into a sequence of indexed in time optimizations solving for the locally optimal at each time transformation and then forward integrating the solution. This is only a locally-in-time optimal method reducing the dimension of the optimization. Assume the deformation fields are generated from velocity fields which are assumed piecewise constant within quantized time increments of size δ, $t_k \doteq k\delta, k = 0, \ldots, K \doteq \frac{T}{\delta}$ giving $v(x,t) = v(x, t_k)$, for $t \in [t_{k-1}, t_k)$, Define the expansion functions only to be a function of $x \in \Omega$, $\psi_k(x)$.

Algorithm 2 (Greedy Image Matching (Christensen [13])) *The sequence of locally optimal transformations $\hat{h}(\cdot, t_k), k = 0, 1, \ldots, K$ is given by $\hat{h}(x, 0) = x$ with*

$$\hat{h}(x, t_k) \quad = \quad \int_{t_{k-1}}^{t_k} -\nabla_x^t \hat{h}(x, \sigma) \hat{v}(x, t_k) d\sigma + \hat{h}(x, t_{k-1}) \text{ with } \hat{v}(x, t_k) \doteq \hat{v}(x, t) \ t \in [t_{k-1}, t_k) \ , \tag{22}$$

$$\text{and } \hat{v}(\cdot, t_k) \quad = \quad \arg \min_{v(\cdot, t_k) = \sum_j \nu_j(t_k) \psi_j} (t_k - t_{k-1}) \int_\Omega \|Lv(x, t_k)\|^2 dx + \frac{1}{\sigma^2} \int_\Omega |I(h(x, t_k)) - I'(x)|^2 dx \ . \tag{23}$$

For each $k = 1, \ldots, K$, the gradient algorithm for local in time updating of Eqns. 22, 23 initializes with $n = 0$, $\nu_j^{(0)}(t_k) = 0$ for all j, and iterates $n = 0, 1, \ldots$:

1. *Calculate transformation for $t \in [t_{k-1}, t_k)$:*

$$h^{(n)}(x, t) \quad = \quad \int_{t_{k-1}}^t -\nabla h^{(n)}(x, \sigma) v^{(n)}(x, t_k) d\sigma + \hat{h}(x, t_{k-1}) \text{ where } v^{(n)}(x, t_k) = \sum_j \nu_j^{(n)}(t_k) \psi_j(x) \ . \tag{24}$$

2. *Calculate gradient perturbation:*

$$\nu_j^{(n+1)}(t_k) \quad = \quad \nu_j^{(n)}(t_k) - \Delta \left((t_k - t_{k-1}) \int_\Omega L^\dagger L v^{(n)}(x, t_k) \psi_j(x) dx + \frac{\partial}{\partial \nu_j} \|h(t_k)I - I'\|^2 \Big|_{\substack{h=h^{(n)}(t_k) \\ v=v^{(n)}(t_k)}} \right) \tag{25}$$

$$\frac{\partial}{\partial \nu_j} \|h(t_k)I - I'\|^2 \quad = \quad \underbrace{\frac{2}{\sigma^2} \int_\Omega (I(h(x, t_k)) - I'(x)) \nabla I(h(x, t_k)) \cdot \frac{\partial h(x, t_k)}{\partial \nu_j} dx}_{\frac{\partial}{\partial \nu_j} D_2(h(t_k))} \ ,$$

and

$$\frac{\partial h(x,t_k)}{\partial \nu_j} = \int_{t_{k-1}}^{t_k} -\nabla_x^t h(x,t))\psi_j(x)dt + \int_{t_{k-1}}^{t_k} -\nabla_x^t \frac{\partial h(x,t)}{\partial \nu_j} v(x,t_k)dt \ .$$

After stopping, define the final iterate as $\hat{v} \doteq v^{(n+1)}$ *given by* $\hat{v}(x,t_k) \doteq \sum_j \hat{\nu}_j^{(n+1)}(t_k)\psi_j(x)$ *with* $\hat{h}(t_k)$ *satisfying Eqn. 22.*

Remark 2 Lattice and eigenfunction implementations. *Christensen et al. [13] has solved the PDE numerically on the lattice using discrete approximations to the differential operators in space. Divide Ω into a discrete lattice Ω^Δ of pixels of fixed size, then $\psi_j(x)$ has it's support on lattice site $x_j \in \Omega^\Delta$. Define L^Δ to be the discretized version of $L = -a\nabla^2 - b\nabla \cdot \nabla + cI$. Then the gradient equations of Eqn. 25 reduce to*

$$\nu_j^{(n+1)}(t_k) = \nu_j^{(n)}(t_k) + \Delta \left((t_k - t_{k-1})L^{\Delta^\dagger}L^\Delta v^{(n)}(x_j,t_k) + \frac{\partial}{\partial \nu_j}\|h(t_k)I - I'\|^2 \Big|_{\substack{h=h^{(n)}(t_k) \\ v=v^{(n)}(t_k)}} \right) \ . \tag{26}$$

3.3 Landmark Deformation in the Lagrangian Frame

The equality constrained exact landmark matching sends $\Sigma \to 0$ giving the equality constrained problem in the limit. For landmark matching, see [16] for the existence of the solution to the limiting problem

$$\hat{v} = \arg\min \|Lv\|^2 \text{ subject to } y_l = h(x_l,T), l=1,\dots,L \ .$$

Here we pose $\Sigma_l, l=1,\dots,L$ non-zero. For small numbers of landmarks, reparameterize the optimization problem following Joshi [16] using the positions of the L-landmarks $\phi_l(\cdot), l=1,\dots,L$ determined by the velocity field $v(x,\cdot)$:

$$\phi(t) \doteq \underbrace{\begin{pmatrix} \phi(y_1,t) \\ \phi(y_2,t) \\ \vdots \\ \phi(y_L,t) \end{pmatrix}}_{3L \times 1} \text{ with } \frac{d\phi(x,t)}{dt} = v(\phi(t,x),t) \ , \ \phi(x,0) = x \ , \ x \in \Omega \ . \tag{27}$$

Following Joshi [16], for landmark matching choose operators of the form $L = diag(-a\nabla^2 + b\nabla\nabla \cdot +cI)$ with $a=1, b=0, c=1$, giving the 3×3 linear matrix diagonal differential operator $L = diag(-\nabla^2 + cI)$ with associated 3×3 covariance matrix smoothing kernel $K(x,y)$ and the $3L \times 3L$ covariance matrix $K(\phi(t))$ playing a fundamental role:

$$K(x,y) \doteq 2\sqrt{c}(2\pi)^{\frac{5}{2}} \begin{pmatrix} e^{-\sqrt{\frac{1}{c}}\|y-x\|} & 0 & 0 \\ 0 & e^{-\sqrt{\frac{1}{c}}\|y-x\|} & 0 \\ 0 & 0 & e^{-\sqrt{\frac{1}{c}}\|y-x\|} \end{pmatrix} \ , \tag{28}$$

$$K(\phi(t)) \doteq \underbrace{\begin{pmatrix} K(\phi(y_1,t),\phi(y_1,t)) & \dots & K(\phi(y_1,t),\phi(y_L,t)) \\ \vdots & \vdots & \vdots \\ K(\phi(y_L,t),\phi(y_1,t)) & \dots & K(\phi(y_L,t),\phi(y_L,t)) \end{pmatrix}}_{3L \times 3L} \ . \tag{29}$$

It will be helpful to define the notation for matrices and block matrices $(K)_{ij}$ to denote the ij entry or block, and for vectors $(A)_i$ to denote the i-th entry.

Theorem 2 (Joshi(1997) [16]) *The optimal diffeomorphism $\hat{\phi}(x,T) = \int_0^T \hat{v}(\hat{\phi}(x,t),\sigma)d\sigma + x$ with $\hat{h} \doteq \hat{\phi}^{-1}$ minimizing Eqn. 8 given by*

$$\hat{v}(\cdot) = \arg\min_v \int_0^T \int_\Omega \|Lv(x,t)\|^2 dxdt + \sum_{l=1}^L [x_l - \phi(y_l,T)]^t \Sigma_l^{-1}[x_l - \phi(y_l,T)] \ , \tag{30}$$

where $\hat{v}$ satisfies

$$\hat{v}(x,t) = \sum_{i=1}^L K(\hat{\phi}(y_i,t),x) \sum_{j=1}^L (K(\phi(t))^{-1})_{ij}\dot{\hat{\phi}}(y_j,t) \text{ and}$$

$$\dot{\hat{\phi}}(y_l,\cdot)_{\substack{l=1,\dots,L}} = \arg\min_{\substack{\phi(y_l,\cdot) \\ l=1,\dots,L}} \int_0^T \sum_{ij} \dot{\phi}(y_i,t)^t (K(\phi(t))^{-1})_{ij}\dot{\phi}(y_j,t)dt + \sum_{l=1}^L [x_l - \phi(y_l,T)]^t \Sigma_l^{-1}[x_l - \phi(y_l,T)] \ ,$$

and $K(\phi(t))$ is the $3L \times 3L$ covariance matrix of Eqn. 29, with inverse $K(\phi(t))^{-1}$ the $3L \times 3L$ matrix with 3×3 block entries $(K(\phi(t))^{-1})_{ij}, i, j = 1, \ldots, L$.

The power of working with the map ϕ is that the optimization has been reduced from velocity fields $v(x, t)$ on $\Omega \times [0, 1]$ to L velocity fields $\dot{\phi}(y_l, t), l = 1, \ldots, L$ on $[0, T]$. For exact landmark matching, care must be taken as $\Sigma \to 0$. This is formulated and proved separately as a corollary in Joshi [16].

The algorithm for landmark matching reduces the problem to a finite dimensional problem by defining the flows on the finite grid of fixed times of size δ, $t_k \doteq k\delta, k = 0, 1, \ldots, K \doteq \frac{T}{\delta}$. Assume velocities piecewise constant within the quantized time intervals, so that for $t \in [t_{k-1}, t_k), \dot{\phi}(y_l, t) = \frac{\phi(y_l, t_k) - \phi(y_l, t_{k-1})}{\delta}$.

Algorithm 3 (Joshi(1997) [16]) *With $\hat{\phi}(y_l, 0) = y_l, l = 1, \ldots, L$, the finite dimensional minimization over $\mathbb{R}^{3LK}$ becomes*

$$
\begin{aligned}
\hat{\phi}(y_l, t_k)_{\substack{l=1,\ldots,L, \\ k=1,\ldots,K}} &= \arg \min_{\substack{\phi(y_l, t_k) \\ l=1,\ldots,L \\ k=1,\ldots,K}} \frac{1}{\delta^2} \sum_{k=1}^{K} \sum_{ij=1}^{L} [\phi(y_i, t_k) - \phi(y_i, t_{k-1})]^t \left(\int_{t_{k-1}}^{t_k} (K(\phi(t))^{-1})_{ij} dt \right) [\phi(y_j, t_k) - \phi(y_j, t_{k-1})] \\
&\quad + \sum_{l=1}^{L} [x_l - \phi(y_l, T)]^t \Sigma_l^{-1} [x_l - \phi(y_l, T)] .
\end{aligned} \tag{31}
$$

The gradient algorithm for minimizing Eqn. 31 initializes with $n = 0$ and $\phi^{(0)}(y_l, t_k) = y_l, l = 1, \ldots L, k = 1, \ldots, K$, and iterates for $n = 0, 1, \ldots$:

1. *Calculate gradient perturbation for each* $\phi(y_l, t_k) \doteq \begin{pmatrix} \phi_1(y_l, t_k) \\ \phi_2(y_l, t_k) \\ \phi_3(y_l, t_k) \end{pmatrix}, l = 1, \ldots, L, k = 1, \ldots, K$:

$$
\phi^{(n+1)}(y_l, t_k) \doteq \begin{pmatrix} \phi_1^{(n+1)}(y_l, t_k) \\ \phi_2^{(n+1)}(y_l, t_k) \\ \phi_3^{(n+1)}(y_l, t_k) \end{pmatrix} = \begin{pmatrix} \phi_1^{(n)}(y_l, t_k) \\ \phi_2^{(n)}(y_l, t_k) \\ \phi_3^{(n)}(y_l, t_k) \end{pmatrix} - \Delta \begin{pmatrix} \frac{\partial}{\partial \phi_1(y_l, t_k)} \left(P(\phi^{(n)}(T)) + D_1(\phi^{(n)}(T)) \right) \\ \frac{\partial}{\partial \phi_2(y_l, t_k)} \left(P(\phi^{(n)}(T)) + D_1(\phi^{(n)}(T)) \right) \\ \frac{\partial}{\partial \phi_3(y_l, t_k)} \left(P(\phi^{(n)}(T)) + D_1(\phi^{(n)}(T)) \right) \end{pmatrix} , \tag{32}
$$

where for $m = 1, 2, 3$,

$$
\begin{aligned}
\frac{\partial}{\partial \phi_m(y_l, t_k)} D_1(\phi(T)) &= 1_T(t_k) \left(2\Sigma_l^{-1}(x_l - \phi(y_l, T)) \right)_m \\
\frac{\partial}{\partial \phi_m(y_l, t_k)} P(\phi(T)) &= 2 \sum_{j=1}^{L} \left(\int_{t_k}^{t_{k+1}} (K(\phi(t))^{-1})_{nj} dt (\phi(y_j, t_k) - \phi(y_j, t_{k+1})) \right)_m \\
&\quad + 2 \sum_{j=1}^{L} \left(\int_{t_{k-1}}^{t_k} (K(\phi(t))^{-1})_{nj} dt (\phi(y_j, t_k) - \phi(y_j, t_{k-1})) \right)_m \\
&\quad + \sum_{j=1}^{L} [\phi(y_j, t_{k+1}) - \phi(y_j, t_k)]^t \frac{\partial \int_{t_k}^{t_{k+1}} (K(\phi(t))^{-1})_{nj} dt}{\partial \phi_m(y_j, t_k)} [\phi(y_j, t_{k+1}) - \phi(y_j, t_k)] ,
\end{aligned}
$$

where

$$
\frac{\partial \int_{t_k}^{t_{k+1}} (K(\phi(t))^{-1})_{nj} dt}{\partial \phi_m(y_j, t_k)} = \int_{t_k}^{t_{k+1}} \left(K(\phi(t))^{-1} \frac{\partial K(\phi(t))}{\partial \phi_m(y_j, t_k)} K(\phi(t))^{-1} \right)_{nj} dt , \tag{33}
$$

and $1_T(t_k) = 1$ for $t_k = T$, and 0 otherwise.

After stopping, define the final iterate as $\hat{\phi} \doteq \phi^{(n+1)}$, and

$$
\dot{\hat{\phi}}(y_l, t) = \frac{\hat{\phi}^{(n+1)}(y_l, t_k) - \hat{\phi}^{(n+1)}(y_l, t_{k-1})}{\delta} , t \in [t_{k-1}, t_k) , k = 1, \ldots, K , \tag{34}
$$

with

$$
\begin{aligned}
\hat{v}(x, t) &= \sum_{l=1}^{L} K(\hat{\phi}(y_l, t), x) \sum_{j=1}^{L} (K(\hat{\phi}(t))^{-1})_{nj} \dot{\hat{\phi}}(y_j, t) , \tag{35} \\
\hat{\phi}(x, T) &= \int_0^T \hat{v}(\hat{\phi}(x, t), t) dt + x , \text{ for all } x \in \Omega .
\end{aligned}
$$

Remark 3 *For choosing initial conditions, for the inexact landmark matching the identity map given by $v(\cdot) = 0$ is used for the initial condition. For exact landmark matching an initial condition is generated from the inexact landmark matching solution following an approach suggested by Younes [121]. Construct an inexact landmark match to within an ϵ-ball of the target landmarks. The initial condition for exact matching is generated by linearly interpolating the inexact landmark match onto the target points. This insures an initial condition which does not cross and maps the landmarks exactly.*

3.4 Brain Mapping

3.4.1 Whole brain matching.

We now show several results from mapping whole macaque brains studied in David Van Essen's laboratory and the human hippocampus with Drs. John Csernansky, Michael Vannier and John Haller. Various such results have been published previously [9, 10, 11, 13, 14, 15, 19, 20, 21, 23]. We begin with an experiment in which a macaque whole brain numbered 87A was taken to be the template, and was mapped into various targets. First various sulcal curves were defined associated with the extremal points of curvature the sulci. Using a dynamic programming algorithm, curves were generated automatically corresponding extremal points of curvature [67, 68] using dynamic programming [70]. The gyri and associated sulci were labeled in several whole brains. The sulcal maps constrain the transformation from one brain to the other. The top left panel of Fig. 2 illustrates several of the sulcal fundus curves which have been identified and placed into the whole brains discretized to 16 points. The deformation field was constrained so that the corresponding points were mapped one-to-one on each other.

These landmarks serve to define the initial low dimensional maps. Such results are shown in Figure 2. The top row of Figure 2 shows a 3-D surface rendering (left column) and of a $640 \times 540 \times 200$ volume from the whole primate brain number 87A which has been analyzed using the hierarchical mapping method. The dark lines are the sulcul maps providing the landmarks. The middle panel shows the result of mapping 87A (left panel) to the target 90C using only the sulcal line constraints to define the transformation. The right panel shows the target brain 90C. The middle row of Figure 2 shows corresponding sections through the template 87A (left) the target 90C (right) and the deformed template (middle). Notice that there is a large difference in the shape and positions of the major subvolumes (the thalamus and the cortical folds) between the undeformed template and the target. Notice the improvement in the alignment of the major subvolumes in the deformed template with the target after the deformation.

The bottom row of Figure 2 shows the final results of the large deformation mapping based on composing image matching and landmark matching. The transformation was parameterized by one displacement vector and velocity vector at each voxel location corresponding to $\approx 20 \times 10^6$ parameters for the large deformation image matching. The transformation of the image volume matches both the target macaque brain surface topography and its internal structure. The generation of the maps allow for meaningful differential geometric features such as volumes and surfaces to be obtained. Notice how they are topologically different when viewed in 2-D, although since the algorithm works completely in 3-D it moves the folds and surfaces around so as to make them match. Examine the third column which is the transformed template. The sections look exquisitely similar to the corresponding target (right column).

3.4.2 Cortical and hippocampus surface matching

Examine the space of two-dimensional smooth surface manifolds $\mathcal{S}$ such as the cortical and hippocampul surface. We map the surfaces by restricting the diffeomorphism on the volume $\Omega \subset I\!\!R^3$ to the embedded submanifolds $S \subset \Omega$. Since the maps $h \in \mathcal{H}$ are diffeomorphisms, their restrictions are diffeomorphisms carrying smooth tangent structures, curvature, etc. Construct the surfaces as triangulated graphs with attached tangent spaces and shape operator associated with locally quadratic charts. The geometry of the quadratic charts transform in the standard way with the Jacobian of the mapping transforming the tangent space, and curvature transformed by the Hessian of the transformation [20]. This forms the basis for the distance between the family of surfaces. For construction of the triangulated graph, contours defining the neuro-anatomically significant substructures of interest are hand traced in each of the sections of the template volume image. From these brains a triangulated graph representing the surface of the substructure is generated using the Marching Cubes algorithm[122]. The top row of Figure 3 shows triangulated graphs representing the cortical and hippocampul surfaces.

S is assumed to be a smooth two-dimensional differentiable sub-manifold of $\Omega \subset I\!\!R^3$, each point in the surface has a neighborhood N diffeomorphic to an open subset D of $I\!\!R^2$, $D \subset I\!\!R^2 \underset{x^{-1}}{\overset{x}{\rightleftharpoons}} N \subset S$, where $x(\cdot)$ is a diffeomorphism at least twice differentiable. The second order local charts (D, x) are established at each point in the surface by locally approximating the surface up to the quadratic terms in a Taylor series. Let $T_{x_i}(S)$ be the tangent space of S at the point x_i with orthonormal span $(b_{x_i}^{(1)}, b_{x_i}^{(2)})$ with unit normal given by the cross-product $n_{x_i} = b^{(1)} \times b^{(2)}$. The quadratic
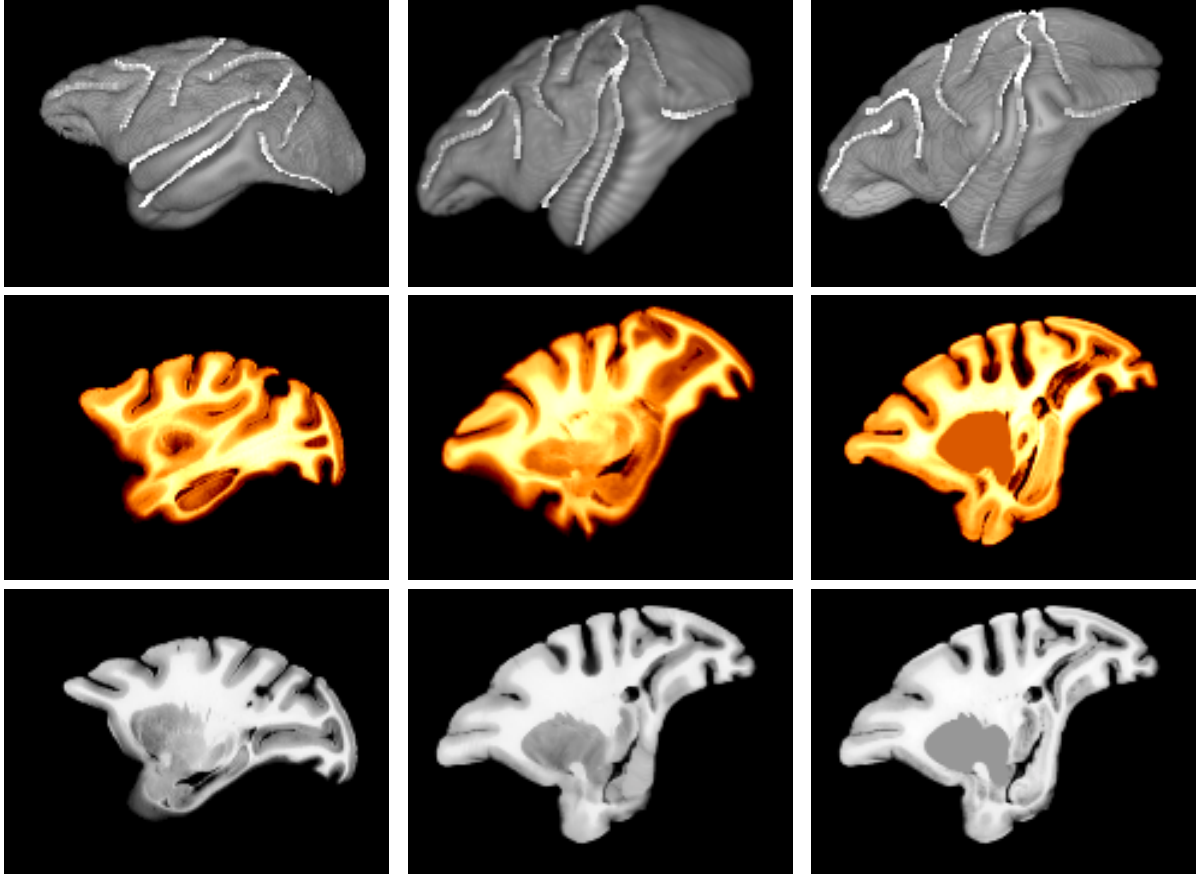
Figure 2: Top row: Left panel shows the volume rendering of the template 87A with sulcal fundus curves depicted superimposed. The middle panel shows the template mapped to the target 90C (middle) using only the sulcal line constraints to define the transformation. The right panel shows the target brain 90C. Middle row: Shows corresponding sections through the template 87A (left) the target 90C (right) and the deformed template (middle) using only the anatomically defined sulcal lines. Bottom row: Shows a sections of the template 87A (left panel) mapped to the target 90C (right panel) using both large deformation landmark matching composed with large deformation image matching. Data were taken from the laboratory of David Van Essen of the Department of Anatomy and Neurobiology of Washington University.

patch approximating the surface passing through the point x_i is written as for $(u, v) \in D$, $x(u, v) \in S$,

$$x(u, v) = x_i + u b_{x_i}^{(1)} + v b_{x_i}^{(2)} + ((u, v) C_i(u, v)^t) n_{x_i} , \tag{36}$$

where C_{x_i} is the symmetric 2×2 matrix of the form $C_{x_i} = \begin{pmatrix} l & m \\ m & n \end{pmatrix}$. The matrix C_{x_i} and tangent space b_1, b_2 are estimated at each point on the surface by a minimum mean squared error algorithm and is based on the scheme described by Hamann[123] following the procedure in Joshi et al. [10].

Once the local charts are established at each vertex, one can study up-to-second-order intrinsic geometrical properties of the surfaces. The maximum and minimum eigenvalues κ_1, κ_2 of the matrix C_{x_i} are the principal curvatures, with the unit vector directions t_1 and t_2 in which these extreme values occur called the principal directions. The principal curvatures describes the maximum and minimum curving degrees of the surface. Shown in the bottom row of Figure 3 is a rendering of a cortical surface in the 87A whole macaque brain (left panel) and a hippocampus surface (right panel). The mean curvature is mapped through a gray color scale. Bright areas represent areas of high positive mean curvature while dark areas represent areas of high negative mean curvature. For construction of the triangulated graph, contours defining the neuro-anatomically significant substructures of interest are hand traced in each of the sections of the template volume image from which a triangulated graph representing the surface of the substructure is generated using the Marching Cubes algorithm[122].
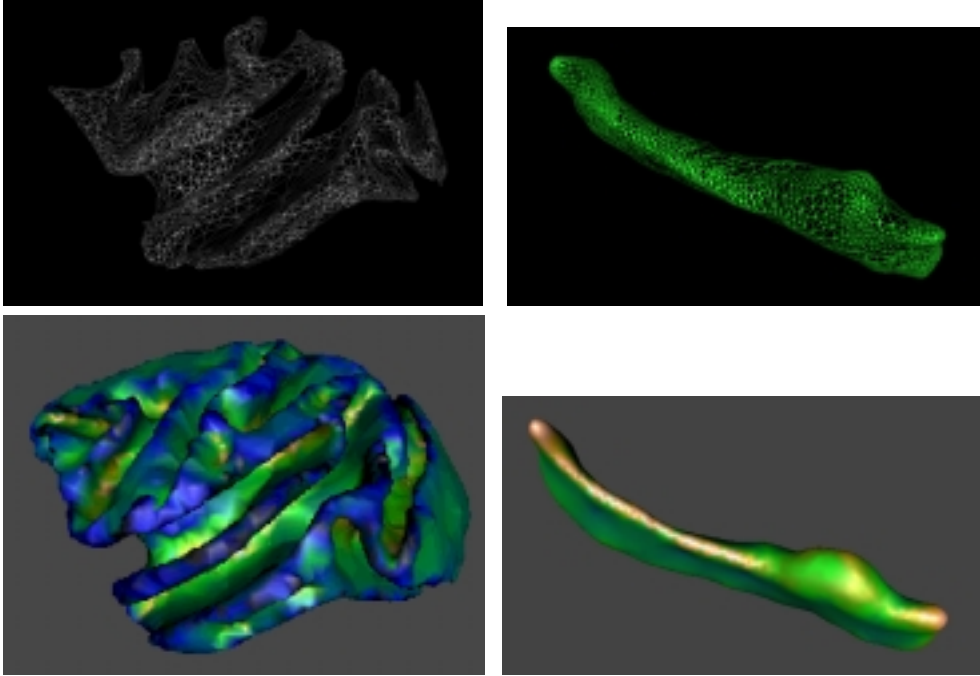
Figure 3: Top row:Left panel shows section of the triangulated graph representing the cortical surface; right panel shows the hippocampus surface. Bottom row: Left panel shows a renderings of the cortical surface with the mean curvature mapped to a gray color scale; right panel shows a rendering of the hippocampal surface with the mean curvature mapped to the same scale. Bright areas represent areas of high positive mean curvature while dark areas represent areas of high negative mean curvature.

Shown in Figure 4 are results obtained from transformation of the neocortical surface of the template 87A into the target whole brain volume 90C. This was accomplished by generating the whole brain maps $h : 87A \rightarrow 90C$, and then restricting the diffeomorphisms to the surfaces $h : S87 \rightarrow S90C$. The template surface is transformed under the diffeomorphic whole brain volume map shown in Figure 2. The whole brain surface was reconstructed in the target by generating a triangulated graph and estimating the C^2 surface using the procedure described in Joshi et al. [10]. The top row shows the resulting quadratic chart which have been surface rendered to exhibit the surface. The bottom row shows a section of the surface embedded in the volume, demonstrating the interface between the white and gray matter.

3.4.3 Hippocampus matching

Figures 5, 6 show similar results for matching hippocampi in human MRI brains. Two anatomies I, I' were selected containing the hippocampus. Figure 5 illustrates the procedure used for hippocampus matching. The top row, left panel shows the template, with the middle panel showing the template mapped through a coarse landmark registration aligning the cube of tissue containing the hippocampus in the target. The right panel shows sections through the target hippocampus cube. The bottom row shows the cube of hippocampus tissue in the template transforming under the large deformation image matching procedure.

The top row, left and middle panels of Figure 6 shows sections through the $128 \times 128 \times 128$ cubes of tissue containing the hippocampus in the template I and target I'. Inserted into both anatomies are the two-dimensional closed surface containing the volume of hippocampus. The intersection with the anatomical volume of the hippocampus surface is depicted via the black line in the anatomical MR section. The bottom row shows the smooth surface mapped in the template via the volume diffeomorphic transformations. The template was mapped onto the target hippocampus; the left panel shows the template surface, the middle panel shows the target surface, the right panel shows the transformed template surface. These surfaces were generated by transforming the template hippocampus through the volume transformations carrying the template onto the targets (see Joshi et al. [20]).
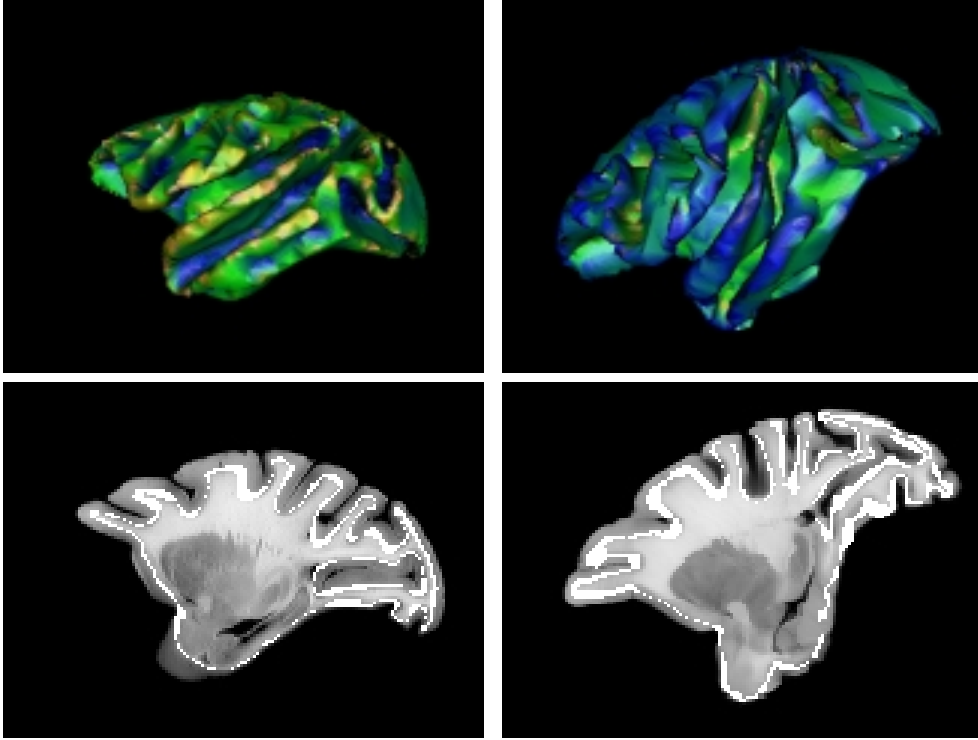
Figure 4: Top row: Left panel shows the C^2 surface in the template 87A; right panel the surface mapped through the volume transformation into 90C. Bottom row: Left panel shows a section with the C^2 surface embedded in the whole macaque template brain; the right panel shows the same section with the template C^2 surface mapped through the transformation to the target.

4   Small Deformations

It would be nice to be able to place the distance measure $\rho(\cdot,\cdot)$, directly on the transformations in $\mathcal{H}$, conjecturing, for example, that $\mathcal{H}$ is a Hilbert space with distance between elements consistent with the norm and inner product of the Hilbert space. However, in general the group $\mathcal{H}$ of diffeomorphisms is curved, addition is replaced by the law of composition. Viewing the elements as vectors in which composition is addition is not valid. This is the fundamental basis behind the transport equation and its introduction in [13, 6, 5]; the kinematic nonlinearity is fundamental to large deformations. Assuming addition of a vector space results in inconsistencies, the resulting distance between elements not necessarily being symmetric since the inverse as a vector may not be a diffeomorphism. For small deformations, these problems approximately disappear and the small deformation solution has an approximate inverse. This will of course only arise in the special setting in which the final time diffeomorphism is virtually identical to an ϵ-scaled version of it's tangent at the identity:

$$h(x) \doteq x - u(x) \approx x - \underbrace{\epsilon v(x,0)}_{u(x)} \ .$$

Then the energy measure may be associated with the ϵ-small deformations, a positive definite quadratic form associated directly with the transformation h itself. In the small deformation setting, this energy approximately corresponds to the one studied above, an energy minimized uniquely at the identity, for example $\|h\|_L^2 = \|L(h-I)\|^2$. As above, assume throughout operators of the form $L = -a\nabla^2 + b\nabla\nabla \cdot +c$ with at least two-derivatives.

4.1   Small Deformation Matching

4.1.1   Landmark matching.

For small deformations, the large deformation matching reduces to the small deformation solution studied by various authors [39, 41, 28, 116, 29, 117, 4, 124]. For exact and inexact landmark matching (see Joshi [16]), choose the 3×3 linear matrix differential operator of the form $L = diag(-\nabla^2 + cI)$. Assume $h(x_l)$ matches to landmark y_l which is a
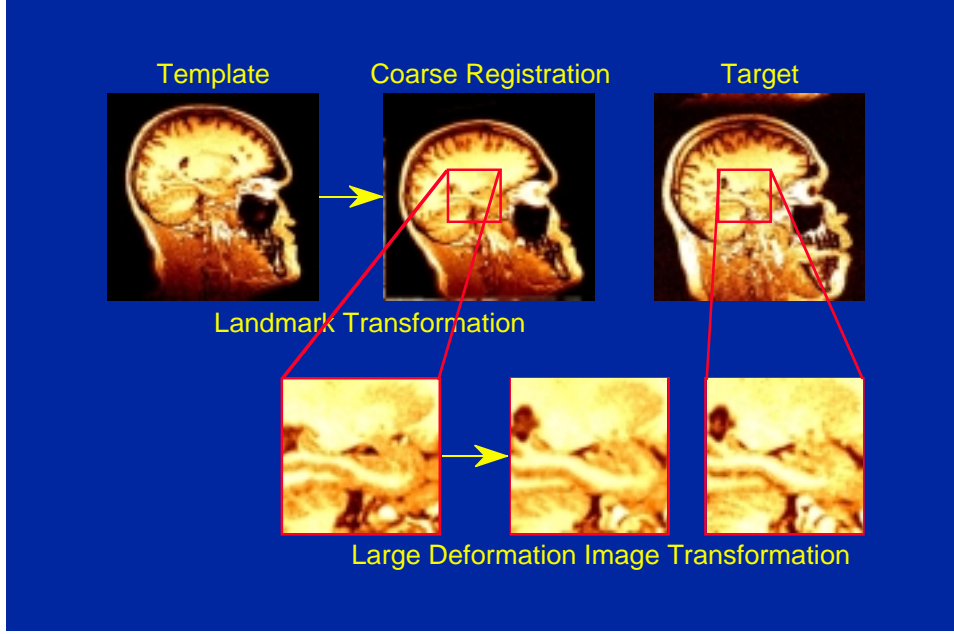
Figure 5: Figure illustrates the procedure used for hippocampus matching. Top row: Left panel shows the template, middle panel shows the template mapped through a coarse landmark registration aligning the cube of hippocampus with the target, right panel shows the hippocampus cube. Bottom row: Left panel shows the template hippocampus, middle panel shows the cube transforming based on image matching after the initial alignment, and right panel shows the target.

small deformation with noise variance Σ_l. Include the affine motions by allowing the template to carry the affine motions

$$(A, a) : y_l \in I\!\!R^3 \mapsto Ay_l + a \in R^3 \ , \ A \in \mathbf{GL}(3), a \in R^3 \ . \tag{37}$$

Following Joshi [9, 16] the small deformation cost is minimized according to $\hat{h}(x) = x - \hat{u}(x)$ with

$$\hat{u}, \hat{A}, \hat{a} \quad = \quad \arg\min_{u,A,a} \int_\Omega \|Lu(x)\|^2 dx + \sum_{l=1}^L [u(x_l) - (x_l - Ay_l - a)]^t \Sigma_l^{-1} [u(x_l) - (x_l - Ay_l - a)] \ . \tag{38}$$

This gives the optimal small deformation

$$\hat{u}(x) \quad = \quad \sum_{l=1}^L \underbrace{\begin{pmatrix} e^{-\sqrt{\frac{1}{c}}\|x_l - x\|} & 0 & 0 \\ 0 & e^{-\sqrt{\frac{1}{c}}\|x_l - x\|} & 0 \\ 0 & 0 & e^{-\sqrt{\frac{1}{c}}\|x_l - x\|} \end{pmatrix}}_{\propto K(x_l, x)} \underbrace{\begin{pmatrix} \hat{\beta}_{l1} \\ \hat{\beta}_{l2} \\ \hat{\beta}_{l3} \end{pmatrix}}_{\beta_l} \tag{39}$$

$$\text{with} \begin{pmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_L \end{pmatrix} \quad = \quad \begin{pmatrix} K(x_1, x_1) + \Sigma_1 & \dots & K(x_1, x_L) \\ & \vdots & \\ K(x_L, x_1) & \dots & K(x_L, x_L) + \Sigma_L \end{pmatrix}^{-1} \begin{pmatrix} x_1 - (\hat{A}y_1 + \hat{b}) \\ \vdots \\ x_L - (\hat{A}y_L + \hat{b}) \end{pmatrix}$$

$$0 \quad = \quad \sum_l \beta_l \ , \ 0 = \sum_l \beta_l y_l^t \ . \tag{40}$$

Changing the operator to $L = \nabla^2$ translates to the problems of Kent [124] with covariance operator proportional to the absolute distance $K(x, y) \propto diag(\|x - y\|)$.
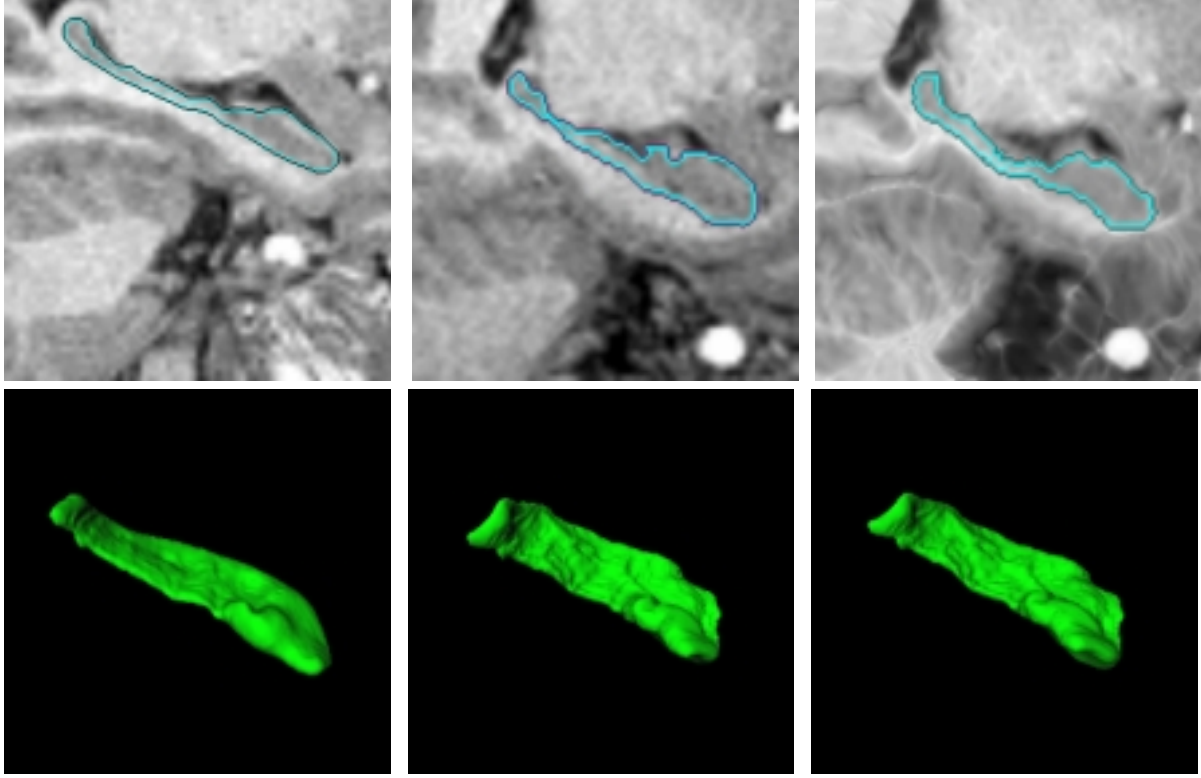
Figure 6: Top row: Left panel shows sections depicting the intersection of the surface in the template volume MRI; middle panel shows the target volume; right panel shows the template transformed. Bottom row: Left panel shows the template surface; middle panel shows the target surface; right panel shows the mapped template surface.

4.1.2 Image Matching

For small deformation image matching, $h(x) \doteq x - u(x), x \in \Omega$ assume $u \in \mathcal{V}$ a Hilbert space with norm $\|u\|_{\mathcal{V}}^2 \doteq \|Lu\|^2$. For image matching, the small deformation problem becomes $\hat{h}(x) = x - \hat{u}(x)$ with

$$\hat{u}(\cdot) = \arg\min_{u \in \mathcal{V}} \int_{\Omega} \|L(x - u(x))\|^2 dx + \int_{\Omega} |I(x - u(x)) - I'(x)|^2 dx \ . \tag{41}$$

Choosing L in the class $L = -a\triangle - b\nabla\nabla \cdot + cI$ gives [4]; for $b = 0$ gives [117].

Algorithm 4 Gradient Algorithm for Small Deformation Image Matching *Initialize with $n = 0$, $\nu^{(n)} = 0$, and iterate $n = 1, 2, \ldots$ calculating and perturbing the transformation:*

$$h^{(n)}(x) \quad \doteq \quad x - u^{(n)}(x) \ , \ u^{(n)}(x) \doteq \sum_{j=0}^{\infty} \nu_j^{(n)} \psi_j(x) \ , \tag{42}$$

$$\nu_j^{(n+1)} \quad = \quad \nu_j^{(n)} - \Delta \left(\int_{\Omega} L^{\dagger} L u^{(n)}(x) \psi_j(x) dx - \frac{2}{\sigma^2} \int_{\Omega} (I_0(h^{(n)}(x)) - I_1(x)) \nabla I_0(h^{(n)}(x)) \cdot \psi_j(x) dx \right) \ . \tag{43}$$

Remark 4 *Reproducing the drifts and gradients in [117, 4], expand $u(\cdot)$ in a complete orthonormal basis $\psi_k(x)$ with $L\psi_k = \lambda_k \psi_k$ eigenfunctions of the space operator L. The gradient in Eqn. 43 reduces to the following form:*

$$\nu_j^{(n+1)} \quad = \quad \nu_j^{(n)} - \Delta \left(|\lambda_j|^2 2\nu_j^{(n)} - \frac{2}{\sigma^2} \int_{\Omega} (I_0(h^{(n)}(x)) - I_1(x)) \nabla I_0(h^{(n)}(x)) \cdot \psi_j(x) dx \right) \ . \tag{44}$$

4.2 Empirical Template Construction in the Small Deformation Setting

For small deformations a template coordinate system can be constructed from averages of transformations. Empirical estimation of the templates for the various subpopulations of interest use the ideas of *minimum mean-squared estimation*(MMSE) for generating the template. Assume a metric distance based on the quadratic energy $\|L(h - I)\|^2$.

Intuitively, the template yet to be discovered should be defined to be the element $I_{temp} \in \mathcal{I}$ which requires the lowest average energy deformation onto the population of anatomies. The template representing the population is defined to be the image $I \in \mathcal{I}$ minimizing the overall energy of the transformation of the population to the template. We now show that for small deformation clusters, a template I_{temp} can be generated by choosing any element of $\mathcal{I}$ which is ϵ close to the cluster, mapping it to the population, and defining the template as the average map applied to I_0.

The important property for ϵ-small deformations is that composition approximately involves moving along the tangent of each transformation; thus the inverse is approximated by moving along the tangent in the opposite direction. For this define ϵ-small deformations, and anatomies which are ϵ-close.

Definition 2 *Given $h_1, \ldots, h_n \subset \mathcal{H}$ a set of diffeomorphisms, then we shall say they are uniformly ϵ-small, or just ϵ-small if for $i = 1, \ldots, n$,*

$$h_i(x) \doteq x - u_i(x) \quad = \quad x - \epsilon \eta_i(x) \ , \ \text{where} \tag{45}$$

$$\sup_{x \in \Omega} \eta_i(x) \quad \leq \quad 1 \ , \ \sup_{x \in \Omega} \|L \eta_i(x)\| \leq K. \tag{46}$$

Given a set of anatomies $I_i, i = 1, \ldots, n$ with diffeomorphisms $h_i \in \mathcal{H}$ carrying one anatomy I to each element of the population $I \underset{h_i^{-1}}{\overset{h_i}{\rightleftarrows}} I_i$, then we shall say the anatomies are uniformly ϵ-close to I if $h_i \doteq I - \underbrace{\epsilon \eta_i}_{\doteq u_i}$ are uniformly ϵ-small.

The importance of these uniformly small deformations is that their compositions and inverses are to an error of order $O(\epsilon^2)$ additive and ϵ small as well.

$$(h_1 \circ h_2)(x) \quad = \quad x - u_1(x) - u_2(x) - u_1(u_2(x))$$

$$= \quad x - u_1(x) - u_2(x) - \epsilon^2 \nabla \eta_1(y)(\eta_2(x) - x) \ , \ y \in [x, u_2(x)] \ ; \tag{47}$$

$$\implies h^{-1}(x) \quad = \quad x + u(x) + O(\epsilon^2) \ . \tag{48}$$

This gives the approximate additive energy and approximate symmetry property required for a quadratic distance function on the transformation to make sense for small deformations:

$$\|h \circ h'\|^2 = \|(h \circ h')^{-1}\|^2 + O(\epsilon^4) = \|L(u + u')\|^2 + O(\epsilon^4) \ . \tag{49}$$

Notice, if there is one element $I \in \mathcal{I}$ which is uniformly ϵ-close to a set of elements, then the population is to error $O(\epsilon^2)$, 2ϵ-uniformly close:

$$h_{nm} \quad = \quad h_m h_n^{-1} = (I - \epsilon \eta_m)(I + \epsilon \eta_n + O(\epsilon^2)) = I - \epsilon(\eta_m - \eta_n) + O(\epsilon^2) \ . \tag{50}$$

Theorem 3 *Given are a set of anatomies $I_n, n = 1, \ldots, N$ of landmarked or imaged brains with the property that there is an $I \in \mathcal{I}$ which is ϵ close to each, $I \underset{\phi_n \doteq h_n^{-1}}{\overset{h_n}{\rightleftarrows}} I_n$.*

Then the small deformation template defined as

$$I_{temp} \doteq \bar{h} I \doteq \{I(\bar{h}(x)), x \in \Omega\} \ , \ \text{where} \ \bar{h} = \frac{1}{N} \sum_{n=1}^{N} h_n \ , \tag{51}$$

has the property that the maps $h'_n, n = 1, \ldots, N$, defining $I_n = h'_n I_{temp}$ are of minimum energy to order $O(\epsilon^4)$.

The template is independent of which ϵ close $I \in \mathcal{I}$ is chosen for the mapping onto the set of N-anatomies.

Proof: Clearly $\bar{h}$ is ϵ-small since

$$\bar{h} \quad = \quad \frac{1}{N} \sum_{n=1}^{N} h_n = I - \epsilon \frac{1}{N} \sum_{n=1}^{N} \eta_n \ , \tag{52}$$

and therefore has inverse $\bar{h}^{-1} = I + \epsilon \frac{1}{N} \sum_{n=1}^{N} \eta_n + O(\epsilon^2)$. To show $I_{temp} \doteq \bar{h} I$ is an image satisfying the minimum energy distance to the family of anatomies, let $I_n = h'_n I_{temp}$ with $I = h' I_{temp}$, and apply the extremum condition subject to the constraints $I_n = \underbrace{h_n \circ h'}_{h'_n} I_{temp}$:

$$\min_{h'} E(h') \quad = \quad \min_{h'} \sum_{n=1}^{N} \|h_n \, h'\|_L^2 = \min_{\eta'} \sum_{n=1}^{N} \|L((I - \epsilon \eta_n)(I - \epsilon \eta') - I)\|^2 \tag{53}$$

$$= \quad \min_{\eta'} \sum_{n=1}^{N} \|\epsilon L(\eta_n + \eta')\|^2 + O(\epsilon^4) \ . \tag{54}$$

The map $\hat{h}' = I + \frac{1}{N}\epsilon\sum_n \eta_n$ attains the minimum to error $O(\epsilon^4)$. Thus $I_{temp} = \bar{h}I$ with $\bar{h} = (h')^{-1} = I - \frac{1}{N}\epsilon\sum_n \eta_n + O(\epsilon^2)$. The construction of the template I_{temp} is independent of sufficiently close $I \in \mathcal{I}$ to the population:

$$
h_i' \quad = \quad h_i \circ (\bar{h})^{-1} = I - \epsilon(\eta_i - \frac{1}{N}\sum_{n=1}^N \eta_n) + O(\epsilon^2) \; , \tag{55}
$$

$$
\Longrightarrow \quad \sum_{i=1}^N E(h_i') = \sum_{i=1}^N \|\epsilon(\eta_i - \frac{1}{N}\sum_n \eta_n)\|^2 + O(\epsilon^4) = \sum_{i=1}^N \|u_i - \frac{1}{N}\sum_{n=1}^N u_n)\|^2 + O(\epsilon^4) \; .
$$

$\blacksquare$

The notion of average maps being connected to small deformations is fundamental. Examine this issue in the setting of a population of landmarked brains in which the minimum energy elements to the populations can be explicitly calculated. Depending upon whether you state the problem in the forward or inverse maps, two different minimum energy elements are arrived at! Define $I_n \underset{h_{nm}=\phi_{nm}^{-1}}{\overset{\phi_{nm}}{\rightleftarrows}} I_m$ satisfying $\phi_n(y_l) = x_{nl}$, $h_n(x_{nl}) = y_l$, $l = 1,\ldots,L$, $n = 1,\ldots,N$. There is the catastrophe of having two different minimun energy templates, which are different for the forward and inverse formulations of the small deformation landmark matching.

Proposition 1 *Examine Eqn. 38 with identity affine motion $A = I, a = 0$. The elements $\hat{I}^h, \hat{I}^\phi \in \mathcal{I} \doteq \Omega^L$ minimizing the inverse and forward energy maps to the population are given by*

$$
\hat{I}^h \quad \doteq \quad \arg\min_{y\in\Omega^L} \min_{h:h(x)=y} \sum_{n=1}^N \|L(h_n - I)\|^2 = (\sum_n K(x_n)^{-1})^{-1} \sum_n K(x_n)^{-1} x_n \; ; \tag{56}
$$

$$
\hat{I}^\phi \quad \doteq \quad \arg\min_{y\in\Omega^L} \min_{\phi:\phi(y)=x} \sum_{n=1}^N \|L(\phi_n - I)\|^2 = \arg\min_{y\in\Omega^L} \sum_n \sum_{ij} (x_{in} - y_i)^t (K(y)^{-1})_{ij}(x_{jn} - y_j) \; . \tag{57}
$$

Proof: Optimizing the inverse and forward mapping fields h_n, ϕ_n gives

$$
h_n(x) - x \quad = \quad \sum_{l=1}^L K(x_{nl},x)\beta_l^n \text{ with } h_n(x_{nj}) = y_{nj} - x_{nj} = \sum_{l=1}^L K(x_{nl},x_{nj})\beta_l^n \; ; \tag{58}
$$

$$
\phi_n(y) - y \quad = \quad \sum_{l=1}^L K(y_l,y)\beta_l^n \text{ with } \phi_n(y_j) - y_j = x_{nj} - y_j = \sum_{l=1}^L K(y_l,y_j)\beta_l^n \; . \tag{59}
$$

Substituting gives the two minimizing templates in terms of minimum quadratic form of Eqns. 56, 57. $\blacksquare$

Remark 5 *For the forward maps, $\Omega \doteq \mathbb{R}^3$, and assume $L = diag(-a\nabla^2 + cI)$ a 3×3 diagonal operator, with 3×3 matrix of the form $K(x,y) = diag(e^{-\alpha|x-y|})$ and define the $3L \times 3L$ matrix $(K(z))$ as in Eqn. 29, with 3×3 matrix entries $(K(z))_{ij} = K(z_i,z_j)$. For the forward maps, $\hat{I}^\phi$ the minimizer solves the nonlinear problem for $\hat{y}_l, l = 1,\ldots,L$,*

$$
0 \quad = \quad \sum_n \sum_j (K(\hat{y})^{-1})_{lj}(x_{nj} - \hat{y}_j) + \sum_n \sum_{ij} (x_{ni} - \hat{y}_i)^t (K(\hat{y})^{-1}\delta_l K(\hat{y})K(\hat{y})^{-1})_{ij}(x_{nj} - \hat{y}_j) \tag{60}
$$

$$
\text{where} \quad \begin{array}{ll} (\delta_l K(y))_{lj} = -\alpha e^{-\alpha|x_l-x_j|} & j \neq l \\ (\delta_l K(y))_{il} = -\alpha e^{-\alpha|x_i-x_l|} & i \neq l \\ (\delta_l K(y))_{ll} = 0 & l = 1,\ldots,L \\ (\delta_l K(y))_{ij} = 0 & i \neq l, j \neq l \end{array} \; . \tag{61}
$$

4.3 Empirically Estimated Templates: Macaque and Hippocampus

To begin select one of the brain volumes $I_0 \in POP - I$; call it the provisory template. All anatomies are mapped diffeomorphically to this one by calculating the set of diffeomorphisms so that $I_0 \underset{h_n^{-1}}{\overset{h_n}{\rightleftarrows}} I_n$, $n = 1,\ldots,N$. Shown in Figure 7 are results of mapping the three brains together to form one template. Panels 1-3 show the three whole brains with the affine group $\mathbf{GL}(3) \otimes \mathbb{R}^3$ removed. The rightmost panel of Figure 7 shows the resulting M.M.S.E. template resulting from the average of two deformation fields. The global scale and skew parameters have been removed. The top row shows the surface rendering, the bottom row sections through the whole brain. Notice how the geometry is exquisitely preserved in the template. The template was generated by applying the empirical average transformation to the provisory template, $\bar{h}I_0$.
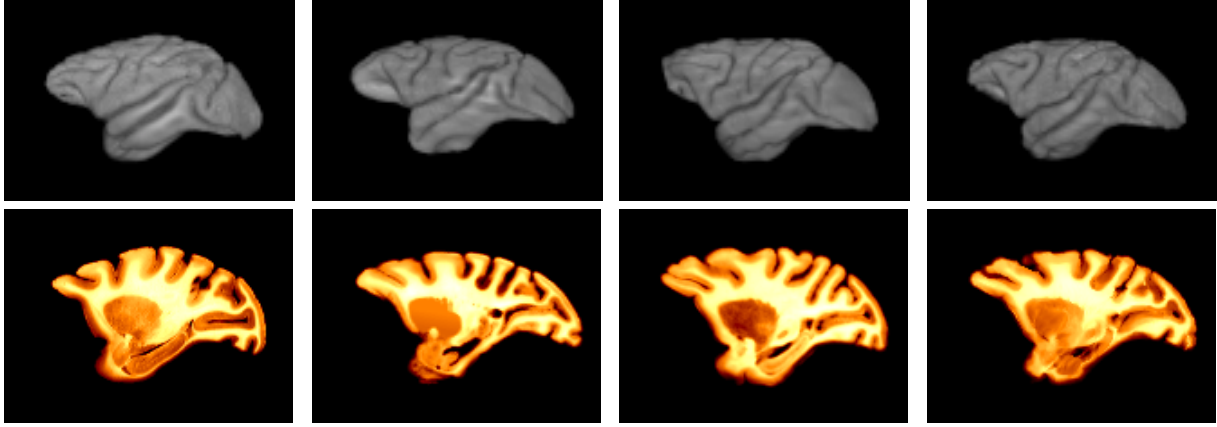
Figure 7: Empirical average of macaque monkey brains, top row showing whole brain volume and bottom row showing sections. Left panel 1 shows 3D volume of 87A, panel 2 shows 3D volume of 90C, panel3 shows 3D volume of 93G, panel 4 shows 3D volume of the empirical average brain.

We have been studying the shape of the human hippocampus [15, 125]. Shown in Figure 8 are mapped hippocampi from three of 30 subjects studied in [125]. The representative image volume cube I_0 was mapped onto 30 target volumes. The resulting transformations were used to transform the template hippocampus surface. Panels 1-3 show three hippocampi which have been mapped from the initial representative surface S_0, so that panels 1-3 show $h_1 S_0, h_2 S_0, h_3 S_0$. Panel 4 shows the result of mapping 30 hippocampi and generating the average map. The composite template surface S_{temp} is generated by mapping the closed surface through the diffeomorphisms onto the populations of target hippocampi, $\{h_n, n = 1, \dots, 30\}$. The representative surface S_0 was then transformed to generate the template:

$$S_{temp} = \bar{h} S_0 \equiv \{\bar{h}(x) \in \Omega : x \in S_0\} \ , \ \bar{h} = \frac{1}{N} \sum_{n=1}^{N} h_n \ . \tag{62}$$

Shown in panel 4 of figure 8 are a template hippocampus generated from three of 30 subjects mapped studied in [15, 125].

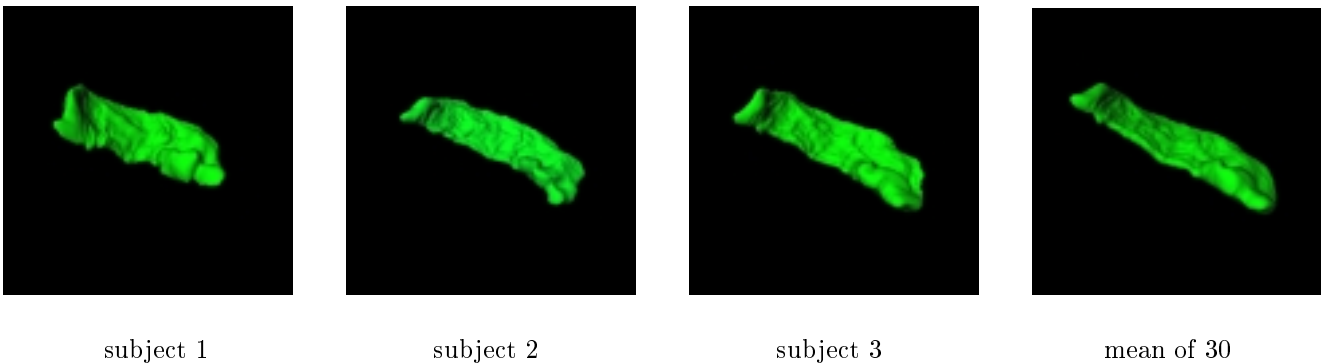subject 1   subject 2   subject 3   mean of 30

Figure 8: Panels 1-3 show maps of the initial hippocampus to three in a population of 30 patients, $h_1 S_0, h_2 S_0, h_3 S_0$ studied in Csernansky et al 1997. Panel 4 shows the composite template $S_{temp} = \frac{1}{30} \sum_{n=1}^{30} h_n S_0$.

5   Empirical Construction of the Prior Distribution

For constructing probability measures of anatomical variation, we characterize these maps as Gaussian processes indexed over the manifolds on which the vector fields are defined. Let's begin in this section with vector fields defined on a smooth connected compact manifold of the full unit cube volume $M \subset \Omega = [0, 1]^3$. Then we concentrate throughout

in this section on the case of the unit cube $M \doteq \Omega \doteq [0,1]^3$. In the unit-cube, we examine spectrum estimation via Autoregressive modelling induced via differential operators from mechanics. To calculate the covariance operator $K : M \times M \to I\!\!R^3$, characterize the vector fields using classical covariance and spectrum estimation of the operators, extensions of ideas well known from the Theory of Stationary Processes [126]. For this *maximum-likelihood estimation* (MLE) is applied directly to estimate from the maps POP the anatomical measures of variability. The covariances are parameterized through the mechanics operators and their polynomial powers. This is for example, as we have done for subcellular organelles in electron microscopy (see [7]).

Specify the vector fields modulo the identity $h = I - u$ as a three-dimensional Gaussian random field $\{U(x), x \in M \subset [0,1]^3\}$ completely specified by its covariance matrix field which is a mapping $K : M \times M \to I\!\!R^3 \times I\!\!R^3$ such that $K(x, x') = K^t(x', x)$ and for any integer n and any n-tuple of 3-vectors $w_1, \ldots, w_n \in I\!\!R^3$ and points $x_1, \ldots, x_n \in \Omega$, $\sum_{i,j=1}^n w_i^t K(x_i, x_j) w_j \geq 0$.

The random field is associated with elements of the Hilbert space [127] with inner product $< f, g >= \int_M f^t(x)g(x)dx$ assumed. Define $\{U(x), x \in \Omega\}$ to be a G.R.F. on the Hilbert space with mean field μ and covariance field K_U if for all $f \in H$, $< f, U >$ is Gaussian distributed with mean $< f, \mu >$, and variance $< f, K_U f >$.

Construct the $\{U(x), x \in M\}$ as a quadratic mean limit using a complete $I\!\!R^3$-valued orthonormal basis $\{\psi_k, k = 0, 1, \ldots\}$, and the U-field given according to

$$U(x) \stackrel{q.m.}{=} \sum_{k=0}^{\infty} U_k \psi_k(x) \ , \tag{63}$$

where U_k are independent Gaussian random variables with fixed means μ_k and variances σ_k^2. The mean and covariance operator of the field becomes

$$\mu(x) = \sum_{k=0}^{\infty} \mu_k \psi_k(x) \ , \quad K_U = \sum_{k=0}^{\infty} \sigma_k^2 \psi_k < \psi_k, \cdot > \tag{64}$$

where $< \psi_k, \cdot >= \int_\Omega \psi_k^*(x) \cdot dx$. Eqn. 63 is meant as the quadratic mean limit minimally requiring the process to have covariance K_U with finite trace so that $\sum_k \sigma_k^2 < \infty$, with mean square integrable, $\sum_k |\mu_k|^2 < \infty$.

The orthogonal expansion reduces the covariance estimation problem to one of estimating the mean and covariance of the orthogonal random variables U_k; $(\mu_k, \sigma_k), k = 1, \ldots$. Since the fields are $I\!\!R^3$-valued it is natural to assume the eigen expansion is real. This is examined below, with the symmetry properties of the expansion understood to reduce it to all real functions.

5.1 Differential and Difference Operators

Clearly, in the above formulation the mean and spectrum $\{\mu_k, \sigma_k^2\}$ and therefore covariances are completely free, corresponding to a non-parametric problem. Since only finite many anatomies can be observed, the class of covariances are restricted using symmetry properties associated with the physical deformation of the tissues. For this, the covariance and random structure is viewed as arising through the fact that $\{U(x), x \in \Omega\}$ is thought to be the solution of a stochastic PDE of the type,

$$L \, U(x) = W(x) \ , \ x \in \Omega \ , \tag{65}$$

$\{W(x), x \in \Omega\}$ a Gaussian random process with covariance K_W.

We will be interested in spectral decomposition, so all of the operators will be normal, induced via difference and differential operators. We shall use the translation operators in $I\!\!R^3$, $(T_h f)(x) = f(x+h)$, to create spatially homogeneous variability operators L, so that $T_h L = L T_h; \forall h \in \Omega$ of the form $L = \sum_{h \in H} A(h) T_h$ with $H \subset \Omega$ a bounded finite set of points, and where the $A(h)$'s are 3×3 normal matrices, so that

$$(Lf)(x) = \sum_{h \in H} A(h) f(x + h)$$

or limits of such finite difference operators in the topology of $L_2(I\!\!R^3, \Omega)$ in which case the domain of L has to be suitably restricted to the proper Sobelev space with norm associated with the continuum operator L. To associate the orthogonal expansion of the Gaussian process, Eqn. 63, with the stochastic PDE of Eqn. 65, choose $\{\phi, \lambda\}$ the eigenelements of the differential operator L according to $L\psi_k(x) = \lambda_k \psi_k(x)$.

Thus far in our work the operators L arise via continuum mechanics construction [119], and correspond to various mixtures of differential operators. We have focussed [13, 6, 5] on the Navier operator $L = -a\triangle - b\nabla \cdot \nabla + cI$.

However, various forms arise, including the bi-harmonic [128] (describing small deformations energetics of thin plates), Laplacian [117]. The general form of the eigenvalues and eigenfunctions assuming general mixtures of differential operators are as follows. Since we will assume cyclo-stationarity of the operators on the unit-cube, they will have eigenfunctions arising from complex exponentials. With $x = (x_1, x_2, x_3) \in \Omega = [0,1]^3$, the complementary variable becomes $\omega_k = (\omega_{k_1}, \omega_{k_2}, \omega_{k_3})$, $\omega_{k_i} = 2\pi k_i$, $i = 1, 2, 3$, and the Fourier basis for periodic functions on $[0,1]^3$ takes the form

$$e^{j<\omega_k, x>} \ , \ <\omega_k, x> = \omega_{k_1} x_1 + \omega_{k_2} x_2 + \omega_{k_3} x_3 \ .$$

As well, since the random U-fields are $I\!R^3$ valued, for each ω_k there corresponds three orthogonal eigenfunctions. It leads to a natural indexing of the eigenfunctions and eigenvalues according to $\{\psi_k^{(d)}, \lambda_k^{(d)}, d = 1, 2, 3\}$. The general form for the eigenfunctions and eigenvalues determines the covariance of the resulting Gaussian process.

Theorem 4 *Let L be a non-singular cyclo-stationary linear normal differential operator on $\Omega = [0,1]^3$ of the form*

$$L = \left(\begin{array}{ccc} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{array} \right) \quad \text{where } A_{il} = \sum_{m=1}^{n_{il}} a_{il}(m) \frac{\partial^{p_{il}(m)}}{\partial x_1^{p_{il}^{(1)}(m)} \partial x_2^{p_{il}^{(2)}(m)} \partial x_3^{p_{il}^{(3)}(m)}} \ , \tag{66}$$

and $p_{il}(m) = \sum_{d=1}^{3} p_{il}^{(d)}(m)$. The eigenfunctions are of the form $\psi_k(x) = \left(\begin{array}{c} c_{k1} e^{j<\omega_k, x>} \\ c_{k2} e^{j<\omega_k, x>} \\ c_{k3} e^{j<\omega_k, x>} \end{array} \right)$ with the vector of constants normalized $\|c_k\|^2 = 1$, and eigenfunctions and eigenvalues for $d = 1, 2, 3$, and $i, l = 1, 2, 3$ satisfying the 3×3 matrix equation:

$$\lambda_k^{(d)} \left(\begin{array}{c} c_{k1}^{(d)} \\ c_{k2}^{(d)} \\ c_{k3}^{(d)} \end{array} \right) = \left(\begin{array}{ccc} \mathcal{A}_{11}(\omega_k) & \mathcal{A}_{12}(\omega_k) & \mathcal{A}_{13}(\omega_k) \\ \mathcal{A}_{21}(\omega_k) & \mathcal{A}_{22}(\omega_k) & \mathcal{A}_{23}(\omega_k) \\ \mathcal{A}_{31}(\omega_k) & \mathcal{A}_{32}(\omega_k) & \mathcal{A}_{33}(\omega_k) \end{array} \right) \left(\begin{array}{c} c_{k1}^{(d)} \\ c_{k2}^{(d)} \\ c_{k3}^{(d)} \end{array} \right) \ , \ \mathcal{A}_{il}(\omega_k) = \sum_{m=1}^{n_{il}} a_{il}(m) \prod_{d=1}^{3} (j\omega_{k_d})^{p_{il}^{(d)}(m)} \ . \tag{67}$$

Proof: Applying $L\psi_k^{(d)} = \lambda_k^{(d)} \psi_k^{(d)}$ gives Eqn. 67. ∎

In general the eigenfunctions are complex. For expansions of real-valued fields, $U_k = U_{-k}^*$, $U_k = <\psi_k, U>$, so that the expansions Eqn. 63 reduce to a real expansion in sines and cosines.

It is natural to perform the computations on the finite lattices with fast Fourier transformation associated with the discrete images. Restrict discussion to the class of discrete operators with periodic boundary conditions. The background space becomes a discrete 3-torus, $\Omega = \{0, 1, \ldots, N-1\}^3$ and the operators are $N \times N \times N$ cyclo-stationary difference operators with addition done modulo N.

The eigenfunctions and eigenvalues are of the same form as in the Theorem, but with the constants given by the discrete Fourier series of the finite difference coefficients.

Corollary 1 *Let the discrete indexed $n = (n_1, n_2, n_3) \in \{0, 1, \ldots, N-1\}^3$ cyclo-stationary operators L be of the form*

$$LU = \left(\begin{array}{c} (LU)_1(n) \\ (LU)_2(n) \\ (LU)_3(n) \end{array} \right) = \left(\begin{array}{c} \sum_{l=1}^{3} \sum_{h \in H} a_{1l}(h) U_l(n+h) \\ \sum_{l=1}^{3} \sum_{h \in H} a_{2l}(h) U_l(n+h) \\ \sum_{l=1}^{3} \sum_{h \in H} a_{3l}(h) U_l(n+h) \end{array} \right) \ , \tag{68}$$

with the operators assumed to have finite support H. The eigenfunctions and eigenvalues $\{\psi_k, \lambda_k\}$ solve the identical matrix equation from Theorem 4, Eqn. 67, with the constants $\mathcal{A}_{il}$ the discrete Fourier transforms of the difference coefficients; for $i, l = 1, 2, 3$, $k \in \{0, 1, \ldots, N-1\}^3$,

$$\mathcal{A}_{il}(\omega_k) = \sum_{h \in H} a_{il}(h) e^{j<\omega_k, h>} \ , \ \omega_k = (\frac{2\pi k_1}{N}, \frac{2\pi k_2}{N}, \frac{2\pi k_2}{N}) \ .$$

5.2 Gaussian Processes and Generalized ARMA Modelling

We shall induce the random $I\!R^3$-valued U-fields via the linear non-singular operator equation $LU = W$, W noise. The eigenfunctions and eigenvalues $\{\psi_k^{(d)}, \lambda_k^{(d)}\}$ derived for the linear differential operators determine the covariance of the resulting Gaussian process.

Let $U_n(x), x \in \Omega$ be a random $I\!R^3$-valued process given by $U_n(x) = \sum_{k=0}^{n} \sum_{d=1}^{3} U_k^{(d)} \psi_k^{(d)}(x)$, the random variables $U_k^{(d)}, k = 1, 2, \ldots$ orthogonal complex Gaussian random variables with mean and variance $EU_k^{(d)} = \mu_k^{(d)}$, $E|U_k^{(d)} - \mu_k^{(d)}|^2 = \frac{|\alpha_k^{(d)}|^2}{|\lambda_k^{(d)}|^2}$. Define L to be a non-singular linear operator having the complete orthonormal basis of eigenelements $L\psi_k^{(d)} = \lambda_k^{(d)} \psi_k^{(d)}$. Then,

Theorem 5 *Let (i) the noise W have covariance $K_W = \sum_k \sum_{d=1}^3 |\alpha_k^{(d)}|^2 \psi_k^{(d)} (\psi_k^{(d)})^\dagger$ and (ii) $\sum_k \sum_{d=1}^3 \frac{|\alpha_k^{(d)}|^2}{|\lambda_k^{(d)}|^2} < \infty$ with $\sum_{k=0}^\infty \sum_{d=1}^3 |\mu_k^{(d)}|^2 < \infty$.*

Then $\{U(x), x \in \Omega = [0,1]^3\}$ is a quadratic mean Gaussian process

$$U(x) \stackrel{q.m.}{=} \sum_{k=0}^\infty \sum_{d=1}^3 U_k^{(d)} \psi_k^{(d)}(x) \quad satisfying \quad LU(x) = W(x) , \tag{69}$$

$$\text{with mean } \mu = \sum_{k=0}^\infty \sum_{d=1}^3 \mu_k^{(d)} \psi_k^{(d)} , \text{ covariance } K_U = \sum_{k=0}^\infty \sum_{d=1}^3 \frac{\alpha_k^{(d)2}}{\lambda_k^{(d)2}} \psi_k^{(d)} (\psi_k^{(d)})^t . \tag{70}$$

If the operator is a difference operator, the Gaussian process has mean and covariance of the form

$$\mu = \sum_{k \in \mathbb{Z}_N^3} \sum_{d=1}^3 \mu_k^{(d)} \psi_k^{(d)} , \ K_U = \sum_{k \in \mathbb{Z}_N^3} \sum_{d=1}^3 \frac{|\alpha_k^{(d)}|^2}{\lambda_k^{(d)2}} \psi_k^{(d)} (\psi_k^{(d)})^\dagger . \tag{71}$$

Proof: That the quadratic mean limit exists follows from the trace class assumption $\sum_k \sum_{d=1}^3 \frac{|\alpha_k^{(d)}|^2}{|\lambda_k^{(d)}|^2} < \infty$ implying that $E\| \sum_{k=n+1}^\infty \sum_{d=1}^3 U_k^{(d)} \psi_k^{(d)} \|^2 \to 0$; call the limit point U. That U satisfies $LU = W$, examine $LU_n - W$ as $n \to \infty$. There are two cases associated with whether W being colored and well defined (trace class) or "white noise".

Case (i): if W is trace class then

$$\sum_k \sum_d |\alpha_k^{(d)}|^2 < \infty \Longrightarrow E\|LU_n - W\|^2 = \sum_{k=n+1}^\infty \sum_{d=1}^3 |\alpha_k^{(d)}|^2 \to 0 \text{ as } n \to \infty . \tag{72}$$

Case (ii): if W is white then $LU = W$ in the generalized distribution sense, and the expansion of $W(x) = \sum_{k=0}^\infty \sum_{d=1}^3 W_k \psi_k(x)$, W_k orthogonal Gaussian random variables, mean 0 and variance 1 only makes sense in integration against smooth test functions (notice K_W is not finite). The left and right hand sides of Eqn. 69 must be equal when integrated against any smooth test function:

$$E| \int f^T(x)(LU_n(x) - W(x))dx|^2 = E| \sum_{k=n+1}^\infty \sum_{d=1}^3 W_k^{(d)} < f, \psi_k^{(d)} > |^2 \tag{73}$$

$$(f - square\ integrable) = \sum_{k=n+1}^\infty \sum_{d=1}^3 | < f, \psi_k^{(d)} > |^2 \to 0 \text{ as } n \to \infty . \tag{74}$$

To show that $\bar{u}, K_U$ are the mean and covariance apply the definition of the mean of a Gaussian field. Take the random variable $< f, U >$ then this is Gaussian distributed with mean $< f, \bar{u} >$ and variance $< f, K_U f >$ for any test function f. That this follows,

$$E < f, U > = \sum_{k=0}^\infty \sum_{d=1}^3 < f, \psi_k^{(d)} > EU_k = \sum_{k=0}^\infty \sum_{d=1}^3 < f, \psi_k^{(d)} > \mu_k \tag{75}$$

proving the first claim. That K_U is the covariance

$$E| < f, U > |^2 = E| \sum_{k=0}^\infty \sum_{d=1}^3 < f, \psi_k^{(d)} > U_k|^2 = \sum_{k=0}^\infty \sum_{d=1}^3 | < f, \psi_k^{(d)} > |^2 \frac{|\alpha_k^{(d)}|^2}{|\lambda_k^{(d)}|^2} \tag{76}$$

proving the second claim. ∎

Remark 6 *It may appear as if the complex Gaussian process has not been completely specified since the correlation between the real and imaginary parts of the expansion coefficients have not been specified, $R(U_k), I(U_k)$, only their total variance. We assume the complex process is circularly symmetric, so that $ER^2(U_k) = EI^2(U_k) = \frac{|\lambda_k|^2}{2}$, and $ER(U_k)I(U_k) = 0$.*

5.3 Empirical estimation of the operator: Generalized ARMA Modelling

The operator L plays the role of a *pre-whitening operator*. Introduce the basic generating operator L_0, and an associated polynomial operator consisting of powers of L_0. This will be a generalization of AR modelling; the operator L becomes the polynomial

$$L = p(L_0) = a_d L_0^d + a_{d-1} L_0^{d-1} \ldots + a_0 I$$

The unknown parameters $a_d, a_{d-1}, \ldots a_0$ estimated from the data. Assume the set $POP - U = \{u_1, \ldots, u_N\}$ of anatomical maps are given, from which the empirical estimates of the expansion coefficients are generated: $u_{kn} = < \psi_k, u_n >$, $n = 1, \ldots, N$. It then follows that the eigenvalues are polynomials of the original eigenvalues.

Corollary 2 *Assume the random U-field satisfies the non-singular operator equation $LU = W$, W a Gaussian process with covariance $K_W = \sum_k \sum_{d=1}^3 |\alpha_k^{(d)}|^2 \psi_k^{(d)} (\psi_k^{(d)})^\dagger$. Let $L = p(L_0) = \sum_{i=0}^d a_i L_0^i$, L_0 a cyclo-stationary operator on $\Omega = [0,1]^3$ with eigenvalues and eigenfunctions $\{\lambda_k^0, \psi_k\}$.*
  Then L has identical eigenfunctions $\{\psi_k\}$ as given by Theorem 4, with eigenvalues satisfying

$$\lambda_k = p(\lambda_k^0) = a_d (\lambda_k^0)^d + a_{d-1}(\lambda_k^0)^{d-1} \cdots + a_0 \ . \tag{77}$$

The variances of the random expansion coefficients are given by $E|U_k^{(d)} - \mu_k^{(d)}|^2 = \frac{|\alpha_k^{(d)}|^2}{|p(\lambda_k^0)|^2}$.

  The maximum-likelihood estimate of the spectrum is given by $\hat{\sigma}_k^{(d)2} = \frac{|\alpha_k^{(d)}|^2}{|\hat{\lambda}_k|^2}$ where $\hat{\lambda}_k = \hat{p}(\lambda_k^0) = \sum_{i=0}^d \hat{a}_i (\lambda_k^0)^i$ and $\{\hat{a}_i, i = 0, 1, \ldots, d\}$ satisfy the MLE equations:

$$\frac{\partial}{\partial a_j} log p(data; a) = 0 \quad \Rightarrow \quad -N \sum_k \frac{(\lambda_k^0)^j}{\hat{p}(\lambda_k^0)} + \sum_k \frac{\hat{p}(\lambda_k^0)(\lambda_k^0)^j}{\alpha_k^{(d)2}} \sum_{n=1}^N |u_{kn} - \mu_k|^2 = 0 \ , \ j = 0, 1, \ldots d \ . \tag{78}$$

We will be estimating parameters of differential operators which are mixtures $L = \sum_i a_i L^{(i)}$. But in this estimation problem L depends upon parameters and to solve the maximum likelihood problem iteratively in general requires re-computation of the eigenvectors. However, if the operators commute, this will not be necessary.

Corollary 3 *Let $L = \sum_i a_i L^{(i)}$, $L^{(i)}$ a normal differential operator of the form given in Theorem 4, with the property that each of the operators commute $L^{(i)}L^{(i')} = L^{(i')}L^{(i)}$. Then $L^{(i)}, i = 1, 2, \ldots$ have identical eigenfunctions $L^{(i)}\psi_k = \lambda_k^{(i)} \psi_k$ for all i. The eigenvalues of $L = \sum_i a_i L^{(i)}$ are $\lambda_k = \sum_i a_i \lambda_k^{(i)}$.*
  The maximum-likelihood estimates satisfy $\hat{\lambda}_k = \sum_i \hat{a}_i \lambda_k^{(i)}$ where $\{\hat{a}_i, i = 1, \ldots\}$ satisfy the MLE equations:

$$\frac{\partial}{\partial a_j} log p(data; a) = 0 \quad \Rightarrow \quad -N \sum_k \frac{\lambda_k^{(j)}}{\left(\sum_i \hat{a}_i \lambda_k^{(i)}\right)} + \sum_k \frac{\left(\sum_i \hat{a}_i \lambda_k^{(i)}\right)\lambda_k^{(j)}}{\alpha_k^{(d)2}} \sum_{n=1}^N |u_{kn} - \mu_k|^2 = 0 \ , \ j = 0, 1, \ldots d \ . \tag{79}$$

Proof: To be proven is that the eigenfunctions are equal for the various operators $L^{(i)}, i = 1, 2, \ldots$. The eigenfunctions take the form $e^{j<\omega_k, x>} \left(c_{k1}^{(i)}, c_{k2}^{(i)}, c_{k3}^{(i)}, \right)^T$, implying that for the operators to have the same eigenfunctions we must show $c_k^{(i)} = c_k^{(i')}$. Since the operators $L^{(i)}, L^{(i')}$ commute, the matrices $(\mathcal{A}^{(i)}), (\mathcal{A}^{(i')})$ from Eqn. 67 commute implying they have the same simple eigenvectors associated with their simple eigenvalues. Thus $c_k^{(i)} = c_k^{(i')}$. ∎

Remark 7 Stationary Navier elasticity operator. *For the elasticity operator [4] $L = -a\nabla^2 + b\nabla \cdot \nabla + cI$, the operator matrix from Theorem 4 has entries $A_{ii} = -a\nabla^2 + b\frac{\partial^2}{\partial x_i^2} + c$, $i = 1, 2, 3$, and $A_{il} = +b\frac{\partial^2}{\partial x_i \partial x_l}$, $i \neq l$, with*

$$A_{ii}(\omega_k) = a\|\omega_k\|^2 - b\omega_{k_i}^2 + c \quad i = 1, 2, 3 \ ; \tag{80}$$
$$A_{il}(\omega_k) = -b\omega_{k_i}\omega_{k_l} \quad i \neq l \ . \tag{81}$$

The eigenelements take the form

$$\begin{aligned}
\psi_k^{(1)}(x) &= \alpha_1(\omega_{k_1}e^{j<\omega_k,x>}, \omega_{k_2}e^{j<\omega_k,x>}, \omega_{k_3}e^{j<\omega_k,x>})^T, \quad \lambda_k^{(1)} = -(2a+b)\|\omega_k\|^2 + c \ , \\
\psi_k^{(2)}(x) &= \alpha_2(-\omega_{k_2}e^{j<\omega_k,x>}, \omega_{k_1}e^{j<\omega_k,x>}, 0)^T \quad , \quad \lambda_k^{(2)} = -a\|\omega_k\|^2 + c \ . \\
\psi_k^{(3)}(x) &= \alpha_3(\omega_{k_1}\omega_{k_3}e^{j<\omega_k,x>}, \omega_{k_2}\omega_{k_3}e^{j<\omega_k,x>}, -(\omega_{k_1}^2 + \omega_{k_2}^2)e^{j<\omega_k,x>})^T \quad , \quad \lambda_k^{(3)} = -a\|\omega_k\|^2 + c \ ,
\end{aligned} \tag{82}$$

with the coefficients α scaling each eigenvector to unit energy

$$\alpha^{(1)} = \sqrt{\frac{1}{\|\omega_k\|^2}}, \alpha^{(2)} = \sqrt{\frac{1}{\omega_{k_1}^2 + \omega_{k_2}^2}}, \alpha^{(3)} = \sqrt{\frac{1}{(\omega_{k_1}^2 + \omega_{k_2}^2)\|\omega_k\|^2}} \ . \tag{83}$$

Since the operator is self-adjoint, $\lambda_k = \lambda_{-k}$, the real eigen elements become $\{\psi_k^{(d)} + \psi_{-k}^{(d)}, 2\lambda_k^{(d)}\}$. Now consider the maximum-likelihood estimation of the parameters (a, b, c) associated with $L = -aL^A + bL^B + cI$, with $L^A u = \triangle u$, $L^B u = \triangledown \cdot \triangledown u$. Examine the method for calculating the eigenvalues. Estimating L (i.e. a,b, and c) from anatomical data depends upon the parameters (a, b, c). To solve the maximization problem iteratively would require re-computation of the eigenfunctions. This is not necessary due to the fact that L^A, L^B, I are normal and commute implying they have identical eigenfunctions. Notice the eigenfunctions are independent of a, b, c as predicted by Corollary 3.

Denote the eigenvalues of A and B by $\{\lambda_k^{(d)A}\}$ and $\{\lambda_k^{(d)B}\}$, then $\lambda_k^{(d)} = (-a\lambda_k^{(d)A} + b\lambda_k^{(d)B} + c)$ since the eigenvalues just add because of the commutativity, Corollary 3. In this case

$$
\begin{aligned}
\lambda_k^{(1)A} &= -(\omega_{k_1}^2 + \omega_{k_2}^2 + \omega_{k_3}^2) \ , \ \lambda_k^{(2)A} = -(\omega_{k_1}^2 + \omega_{k_2}^2 + \omega_{k_3}^2) \ , \ \lambda_k^{(3)A} = -(\omega_{k_1}^2 + \omega_{k_2}^2 + \omega_{k_3}^2) \ , \\
\lambda_k^{(1)B} &= -(\omega_{k_1}^2 + \omega_{k_2}^2 + \omega_{k_3}^2) \ , \ \lambda_k^{(2)B} = 0 \ , \ \lambda_k^{(3)B} = 0 \\
\lambda_k^{(1)C} &= 1 \ , \ \lambda_k^{(2)C} = 1 \ , \ \lambda_k^{(3)C} = 1 \ .
\end{aligned}
$$

Hence the eigenvalues only have to be computed once, even though the coefficients a, b, c change during the non-linear estimation procedure.

5.4 What Spectral Representations Can Be Reached?

From the above, we have a method for estimating the cyclo-stationary parametrically defined spectra generated from a primitive operator L_0. But the question arises, *how general a family of operators can be obtained from L_0?*, or equivalently, *how general a family of cyclo-stationary spectra?*. To illuminate, restrict to the class $\mathcal{L}$ of discrete self-adjoint operators with periodic boundary conditions, $\Omega = \mathbb{Z}_N^3$ the discrete 3-torus and the operators as $N \times N \times N$ arrays of circulant type. These operators all commute and have the same eigen-vectors $\psi_k = \psi_{k_1 k_2 k_3}$ with the eigenvalues $\lambda_k, k = (k_1, k_2, k_3) \in \mathbb{Z}_N^3$ of L_0 given by corollary 1. Some may be multiple, which will be the case typically if the operator has symmetry properties. Let the corresponding multiplicities be

$$m_k = m_{k_1 k_2 k_3}; k_1, k_2, k_3 = 1, 2, \ldots N \ .$$

For this, define the operators generated from from polynomials of L_0 as $\mathcal{L}(L_0) \subset \mathcal{L}$:

$$\mathcal{L}(L_0) \doteq \{L = p(L_0) : p(A) = \sum_{i=0}^{d} a_i A^i\} \ . \tag{84}$$

We will show that if an operator has lower multiplicites than the generating operator L_0 it cannot be obtained from L_0. For example, if L_0 has the eigen-values $(2, 2, 3, 3, 4, 4, 5, 5)$ so that the multiplicities are 2, then we can get as an example $L \in \mathcal{L}$ with the eigenvalues $(1, 1, 5, 5, 4, 4, 3, 3)$ or $(1, 1, 1, 1, 7, 7, 6, 6)$, but not one with eigenvalues $(1, 2, 9, 9, 8, 8, 7, 7)$.

Theorem 6 *An operator $L \in \mathcal{L}$ can be obtained as a polynomial in L_0, $L \in \mathcal{L}(L_0)$, if its multiplicities are not less than those of L_0:*

$$m_k \geq m_k^0 \ .$$

Proof: Write the operators L_0 and L in spectral decomposition

$$L_0 = \sum_k \lambda_k P_k \ , \ L = \sum_k \lambda_k^0 P_k \ , \tag{85}$$

where P_k is the projection operator down to the sub-space spanned by ψ_k:

$$P_k = \psi_k < \psi_k, \cdot > : (\mathbb{R}^3)^{\mathbb{Z}_N^3} \to span\{\psi_k\} \ . \tag{86}$$

Notice, operators being normal implies eigenfunctions are orthogonal. For any polynomial $p(\cdot)$ transforming L_0 to L, $L = p(L_0)$, we obtain the transformed eigenvalues $\lambda_k^L = p(\lambda_k^0)$. It remains to show that there is a polynomial that makes the transformed eigenvalues equal to the ones prescribed for L: $p(\lambda_k^0) = \lambda_k^L, k = 1, \ldots N^3$. We have two finite

sets $\Lambda^0 = \{\lambda_k^0\}, \Lambda^L = \{\lambda_k^L\}$ of numbers, the *distinct* eigen-values of L_0 and L respectively and we can therefore find a polynomial of some degree d such that the numbers in the first set are mapped into the numbers in the second. The only case when this is not possible is when $|\Lambda| > |\Lambda^0|$ but this can be ruled out if the multiplicities behave as required in the proposition. ∎

Essentially, this theorem is a statement of operator symmetry. The new variability operator L should not have less symmetry structure than the original one L_0. It is therefore not advisable to choose as first approximation an L_0 operator with a lot of symmetry. This answers definitively how general a class of auto-regressive spectra can be obtained.

Remark 8 1-D and 3-D Laplacian examples. *It is informative to examine two examples, 1-dimensional and 3-dimensional cases. In 1-D consider the second difference operator*

$$L_0 u(n) = u(n+1) - 2u(n) + u(n-1); n \in Z\!\!\!Z_N .$$

The eigen-values are then

$$\lambda_k^0 = 2(cos\frac{2\pi kn}{N} - 1); k = 0, 1 \ldots N-1 .$$

Say that N is even; the odd case is dealt with in the same way. Then the multiplicities $m_k^0 = 2$, for all k so that if $L \in \mathcal{L}$ is to be expressed as a polynomial in L_0 it must have double eigen-values. But the eigen-values of any symmetric circulant matrix $L = (c_{n-m}; n, m = 1, 2 \ldots N)$ are proportional to $\lambda_k \propto \sum_n c_n exp^{\frac{j2\pi kn}{N}}$. Since the matrix is symmetric, $c_{n-m} = c_{n-m}$, the c-sequence is even and the eigen-values appear in pairs. Hence L can be expressed as a polynomial and to use the class $\mathcal{L}$ implies no loss of generality.

Examine the 3-D discrete Laplacian $L_0 = \Delta$ which has the eigen-values

$$\lambda_k^0 = 2(\cos\frac{2\pi k_1}{N} + \cos\frac{2\pi k_2}{N} + \cos\frac{2\pi k_3}{N} - 6) .$$

This operator has more symmetry. Not only is there the symmetry $\lambda_{-k_1 k_2 k_3}^0 = \lambda_{k_1 k_2 k_3}^0$, but as well the symmetries $\lambda_{k_2 k_1 k_3}^0 = \lambda_{k_1 k_2 k_3}^0$, and so on. In this case the restriction to the class $\mathcal{L}$ is essential.

5.5 Probabilistic Representations of Submanifold Surfaces

The beauty of the general definition of $\{U(x), x \in \Omega\}$ as a Gaussian random field through it's inner product with functions in the Hilbert space is that it immediately generalized to more general background spaces such as surface submanifolds $S \subset I\!\!R^3$. We are interested in 2-dimensional manifolds associated with embedded surfaces such as the hippocampus. We have been involved in quantifying the variation of the shape of the hippocampus subvolumes in brains via magnetic resonance imagery [55, 56, 14, 15] for inference of neuropsychiatric disease. To quantify the shape of the hippocampus, we use complete orthonormal bases representing the normal deviations of populations of hippocampi.

To construct the variability representation, define the primary starting hippocampal surface to be a 2-dimensional closed smooth manifold S. Then a Gaussian random field $U(x), x \in S$ is specified by its covariance matrix field, mapping $K : S \times S \to I\!\!R^3 \times I\!\!R^3$. Notice this is a Gaussian field on the submanifold S. The Gaussian field $U(x), x \in S$ is constructed using a complete orthonormal basis defined on the template hippocampus S, so that the basis expands all square integrable vector fields on S. The U-field becomes $U(x) = \sum_k U_k \psi_k(x)$, where U_k are independent Gaussian random variables with fixed means μ_k and variances σ_k^2. The mean and covariance operator of the field

$$\mu = \sum_k \mu_k \psi_k , \ \ K_U = \sum_{k=1}^\infty \sigma_k^2 \psi_k \psi_k^\dagger . \tag{87}$$

The means and variances are estimated from the family of maps.

The maximum-likelihood estimates for the mean and variances of the orthogonal Gaussian random variables $U_k \sim N(\mu_k, \sigma_k^2)$ given the set of anatomical maps $h_n(x) \doteq x - u_n(x), n = 1, \ldots, N$ from a Gaussian random field with mean-field and covariance (μ, K_U) given by Eqn. 64 becomes

$$\hat{\mu}_k = \frac{1}{N} \sum_{n=1}^N < \psi_k, u_n > , \ \hat{\sigma}_k^2 = \frac{1}{N} \sum_{n=1}^N | < \psi_k, u_n > -\hat{\mu}_k|^2 . \tag{88}$$

To generate complete orthonormal bases on a C^2 surface M we use finite element codes for thin shells analogous to that done in the non-statistical setting in [49] as well as principle components methods. For the finite elements methods, the basis and variances are chosen to correspond to differential operators associated with thin elastic shells (see [129] for details). The differential operator derived from the linear shell theory of elasticity is discretized into a difference

operator. The eigen vectors and the eigen values $\{(\psi_k, \lambda_k)\}$ of the difference operator are computed numerically and used as the orthonormal basis. To model the elastic surface and compute the corresponding eigen vectors and eigen values $\{(\psi_k, \lambda_k)\}$ we use the finite element package ADINA (see [130] for details). This approach allows for the generation of the complete orthonormal basis of "Surface Harmonics" for any smooth geometry. The familiar example are the spherical harmonics which are the eigen functions of the Laplacian operator on the surface of the sphere. Shown in the left panel of Figure 9 are are four surface harmonics for the hippocampus. All surface harmonics were generated using the software package ADINA. A template hippocampus was generated with smooth geometry, from which ADINA numerically calculated the small deformation energetics corresponding to thin elastic shells. The lowest order surface harmonics are always the linearized rigid-body modes of translation and rotation and are in the null space of the elastic operator. Herein we concentrate on our principle components approach. Begin by defining the composite template surface S_{temp}. Associate with the set of imagery $I_n, n = 1, \ldots, N$ the surfaces $S_n, n = 1, \ldots, N$ of two-dimensional C^2 closed surfaces. Choose representative S_0 as the provisory C^2 surface and constructing on it a triangulated graph along with a family of compatible locally quadratic charts at every point in the graph. This is depicted in Figure 5 showing the starting hippocampus. The representative anatomy I_0 is mapped to the family $I_0 \underset{h_1^{-1}}{\overset{h_{n}}{\rightleftarrows}} I_n$. The composite template surface was generated by mapping the closed surface S_0 through the diffeomorphisms onto the family of target surfaces according to $S_0 \underset{h_1^{-1}}{\overset{h_{n}}{\rightleftarrows}} S_n$. The template becomes

$$S_{temp} = \bar{h}S_0 \doteq \{\bar{h}(x) \in \Omega : x \in S_0\} \ , \ \bar{h} = \frac{1}{N}\sum_{n=1}^{N} h_n \ . \tag{89}$$

Such a template generated from 30 average maps is depicted in panel 4 of Figure 8.
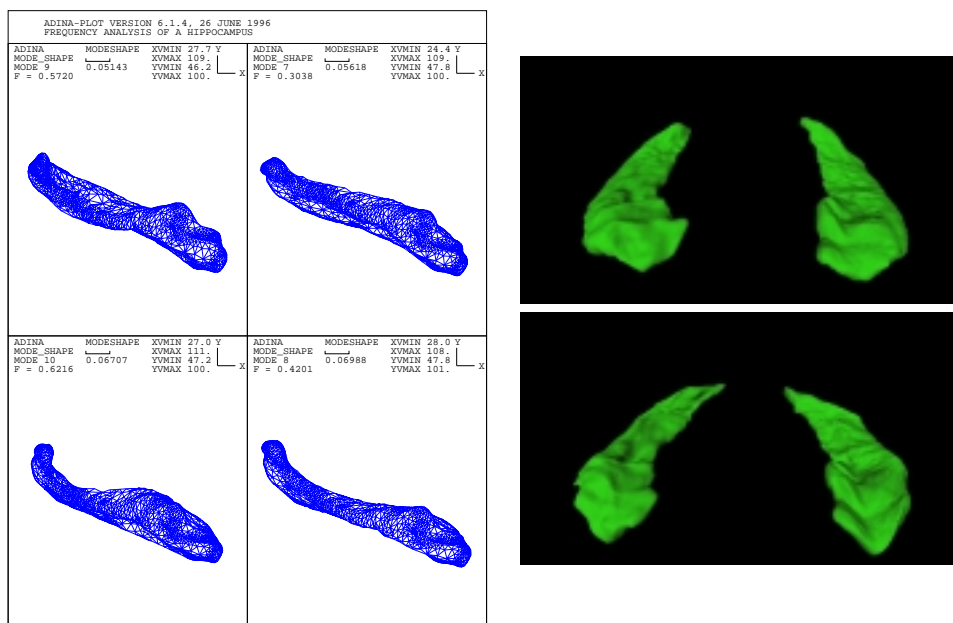

Figure 9: Left column shows the first four surface harmonics associated with the surface of the mean Hippocampus. The surface harmonics are computed using the ADINA package. Right column shows the first two eigen shapes of the left and right hippocampus generated from a population of maps of normals and schizophrenics. Taken from Joshi,1997 Phd. thesis

To construct the variability representation, define the Gaussian random field $U(x), x \in S$ on the domain S is completely specified by its covariance matrix field, mapping $K : S \times S \rightarrow \mathbb{R}^3 \times \mathbb{R}^3$. Notice this is a Gaussian field on the submanifold S. The right panel of Figure 9 shows the first two eigenshapes generated from the empirical covariances of the population of ten maps onto a population of ten anatomies.

To illustrate the variability present in the population, shown in Figure 10 are randomly generated hippocampi mixed with actual hippocampi. These were generated by drawing Gaussian random variables for expanding the orthogonal expansion, and then generating the transformation field applied to the template starting hippocampus.
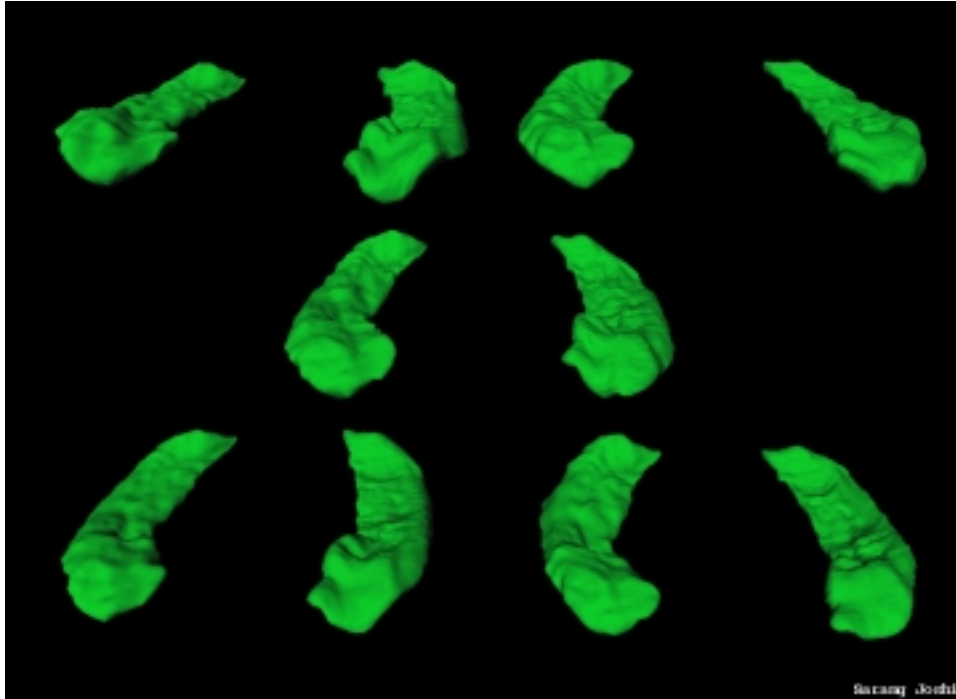
Figure 10: Figure shows real and randomly generated hippocampi generated from the empirically constructed distribution of 30 hippocampus maps. Taken from Joshi Ph.D. thesis.

Wang and Joshi et al. have found that scale and volume are not powerful discriminants of group difference in the two populations; however, shape difference is. Shown in the top row of Figure 11 are the changes in mean hippocampal shape seen in the right hippocampus of 30 subjects, 15 normal and 15 schizophrenic. Left panel shows difference between the mean hippocampus between 15 schizophrenics and the 15 normals average hippocampus. Notice that the superior and lateral aspects of the hippocampal head are deformed inwards in the schizophrenic group. This is consistent with that found as well on the left side by Csernansky, Wang and others. To illustrate that the shape change is relatively highly localized, the right panel shows the variability of 15 normal hippocampus variations around the mean. The color display was generated by expanding the vector fields in the identical 6 eigenfunctions which are used to generate the p-value in the test, and plotting the variability of the field as a function of position on the hippocampus.

Based on these fields, the optimal test statistics were computed for discriminating between the two populations. As shown by Joshi, testing based on scale and volume gives no significant discrimination, $p = .19$ and $p = .27$. The bottom row of Figure 11 shows the jack-knifed log-likelihood ratio tests based on scale (left panel) and shape (right panel). Each test had 29 fields from which the empirical statistics means and covariances were calculated forming the optimal Bayes test; the remaining field was then tested against these empircal means and covariances. This was repeated sequentially 30 times. Notice, the lack of separation for the scale test (left panel). In contract, notice the distinctive separation and clustering of the shape statistics based on log-likelihood ratios (right panel). Use of the first six eigenvectors showed a statistically significant different between the two groups with an F=2.68,p=0.040. [16].

Following Joshi [16], we have also examined distribution free testing on shape variation between the populations. The above mentioned p-values were derived based on the assumptions that the populations *Control* and *Schizophrenics* was were Gaussian distributed with different means and common covariance. We have used Fisher's method of randomization to derive a distribution free estimate of the level of significance of the difference. In Joshi [16] we estimated empirically the probability distribution via simulation from which the significant p-value was computed. In the distribution free testing, we similarly found significant p-values (0.02) demonstrating a statistically significant difference in the populations.
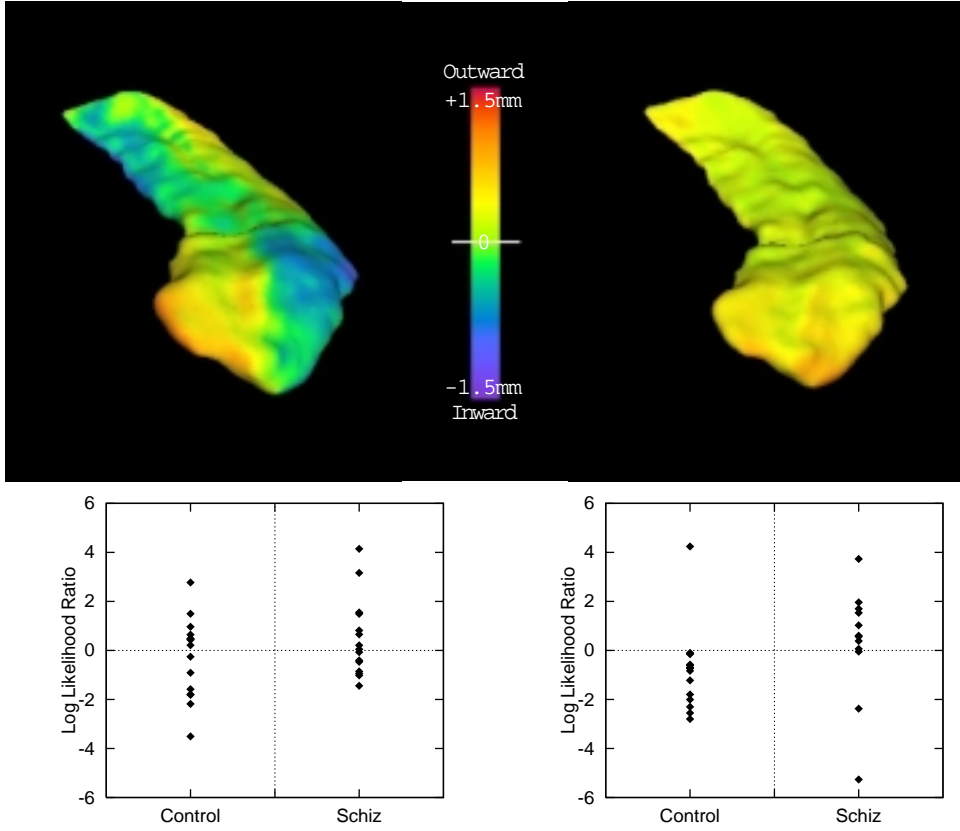
Figure 11: Top row: Left panel shows difference between the mean hippocampus between 15 schizophrenics and the 15 normals average hippocampus. Right panel shows the variability of 15 normal hippocampus variations around the mean expanding the vector fields in the 6 eigenfunctions which generate the p-value. Bottom row: Shows the Bayesian classification log-likelihood ratio for the 15 controls and 15 normals jack-knifed sequentially 30 times. Left panel shows the scale test, right panel the shape test based on 6 eigenfunctions. Taken from Joshi,1997 PhD. thesis.

6 Large Deformation Prior Measures

6.1 Large Deformation Prior Distribution

The large deformation prior is most naturally placed on the random velocity fields and its associated Hilbert space structure. The prior is induced on the space of diffeomorphisms by constructing transformations from flows corresponding to $\mathbb{R}^3$ valued Gaussian random velocity vector fields. The velocity fields are modelled as Gaussian random vector fields $V(\cdot)$ on $\Omega \times [0,1] \doteq [0,1]^4 \subset \mathbb{R}^4$. If the velocity vector fields are sufficiently smooth then the solution to the ODE exists and defines a unique flow of diffeomorphisms. Smoothness will come as above by assuming the velocity fields are solutions of stochastic partial differential equations of the type $L\,V = W$, L a self-adjoint positive differential operator. To simplify, L will differentiate in time and space giving smoothness in both variables, and we will assume diagonal operators with the three components of the noise process $W^{(1)}, W^{(2)}, W^{(3)}$ assumed independent and identically distributed. This implies the three components of the velocity vector field $V^{(1)}, V^{(2)}, V^{(3)}$ are also independent. This is for convenience, and all results hold more generally.

It is convenient to index the single variable $x \in [0,1]^4$. We now study the sample path properties of scalar Gaussian random fields $V(x) \in \mathbb{R}^1, x \in [0,1]^4$ of the type defined above where L is a constant coefficient differential operator of the form

$$L = \sum_{m=1}^{n} a(m) \frac{\partial^{p(m)}}{\partial x_1^{p_1(m)} \cdots \partial x_4^{p_4(m)}} + \epsilon \;, \quad \text{with} \; \sum_{l=1}^{4} p_l(m) = p(m) \;. \tag{90}$$

The eigenvalues and eigenfunctions $L\psi_k(x) = \lambda_k \psi_k(x)$ are of the form

$$\psi_k(x) = e^{j2\pi \langle k, x \rangle} \;, \lambda_k = \sum_{m=1}^{n} a(m) \sum_{l=1}^{4} (j2\pi k_l)^{p_l(m)} + \epsilon \;, \; k \in \mathbb{Z}^4 \;. \tag{91}$$

Also assume that in general the operator L is *strongly elliptic* of order q (Reed and Simon [131]) where for all but a finite k

$$Re\{\lambda_k\} = Re\{\sum_{m=1}^{n} a(m) \sum_{l=1}^{4} (j2\pi k_l)^{p_l(m)} + \epsilon\} \geq \|k\|^{2q} .$$

Remark 9 *The q power of the Laplacian, $L = (-\nabla^2 + I)^q$ has such a property, since with*

$$\nabla^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} + \frac{\partial^2}{\partial x_4^2} , \tag{92}$$

then the eigenvalues are $\lambda_k = \|2\pi k\|^2 + 1$; for L^q, $\lambda_k = (\|2\pi k\|^2 + 1)^q$.

Theorem 7 (Joshi [16]) *Let $V(x), x \doteq (x_1, x_2, x_3, x_4) \in \Omega \doteq [0,1]^4$ be a real-valued Gaussian random field defined in the quadratic mean sense as above to satisfy the partial differential equation*

$$LV(x) = W(x) , x \in [0,1]^4 ,$$

where L is strongly elliptic of order q and the noise process $W(x)$ is assumed to have a spectral representation with spectral variances variances $|\alpha_k|^2$ bounded above by $\|k\|^p$ except on a finite set $F \subset \mathbb{Z}^4$. If $4q - 4 - 2p > 4$ then $V(x)$ has sample paths which are almost surely continuously differentiable.

Proof: To prove sample path continuity we use the Kolmogorov criteria following Amit and Piccioni [132]; then $V(x), x \in \mathbb{R}^4$ is *a.s.* sample path continuous if and only if there exists constants $\alpha, \beta, c > 0$ such that

$$E|V(x) - V(y)|^\alpha \leq c\|x - y\|^{4+\beta} .$$

Since V satisfies the stochastic partial differential equation $L V(x) = W(x)$, $x \in [0,1]^4$, use the spectral representation giving

$$
\begin{aligned}
E|V(x) - V(y)|^2 &= \sum_{k \in Z^4} \frac{|\alpha_k|^2}{|\lambda_k|^2} |e^{j2\pi \langle k,x \rangle} - e^{j2\pi \langle k,y \rangle}|^2 = \sum_{k \in Z^4} \frac{|\alpha_k|^2}{|\lambda_k|^2} (2 - 2\cos 2\pi \langle k, x - y \rangle) \\
&= \sum_{k \in Z^4} \frac{|\alpha_k|^2}{|\lambda_k|^2} (4\sin^2 \pi \langle k, x - y \rangle) \leq \sum_{k \in Z^4} \frac{|\alpha_k|^2}{|\lambda_k|^2} c|\langle k, x - y \rangle|^2 \\
&\leq \sum_{k \in Z^4} |\frac{|\alpha_k|^2}{|\lambda_k|^2} c\|x - y\|^2 \|k\|^2 \overset{(a)}{\leq} c\|x - y\|^2 (c_1 + \sum_{k \in \mathbb{Z}^4/F} \frac{\|k\|^{2(p+1)}}{\|k\|^{4q}}) ,
\end{aligned}
\tag{93}
$$

where (a) uses the fact that L is strongly elliptical of order q.

To establish the Kolmogorof conditions with $\alpha = 6, \beta = 2$, since $V(x)$ is a zero mean Gaussian field, then $V(x) - V(y)$ is a zero mean Gaussian random variable with the moment generating function given by $\Psi(s) = e^{\frac{\sigma^2 s^2}{2}} = \sum_{l=1}^{\infty} \frac{\sigma^{2l} s^{2l}}{2^l l!}$, where $\sigma^2 = E|V(x) - V(y)|^2$. Then for positive integer $l = 3$

$$
\begin{aligned}
E|V(x) - V(y)|^{2l} &= \frac{2l!}{2^l l!} (E|V(x) - V(y)|^2)^l \\
&\leq c(\|x - y\|^2)^l = c\|x - y\|^{4+2} \text{ for } l = 3 .
\end{aligned}
\tag{94}
$$

That the process has a.s. sample path continuity as the series converges for $4q - 2 - 2p > 4$.

To show it has a.s. continuously differentiable sample paths, examine the components of the derivative of the vector field $V'(x)$ with mean square derivative, $V'(x)$ having spectral representation given by

$$V'(x) \doteq \underbrace{(V_1'(x), V_2'(x), V_3'(x), V_4'(x))}_{(\frac{\partial V(x)}{\partial x_1}, \cdots, \frac{\partial V(x)}{\partial x_4})} = \sum_{k \in Z^4} \frac{j2\pi k W_k(\omega)}{\lambda_k} e^{j2\pi \langle k,x \rangle} .$$

As $V'(x)$ is a Gaussian random vector field

$$
\begin{aligned}
E\|V'(x) - V'(y)\|^{2l} &\leq k^l \sum_{j=1}^{4} E(V_j'(x) - V_j'(y))^{2l} \\
&= k^l \frac{2l!}{2^l l!} \sum_{j=1}^{4} (E(V_j'(x) - V_j'(y))^2)^l .
\end{aligned}
\tag{95}
$$

32

It is enough to show that for each j, $E(V_j'(x) - V_j'(y))^2 \leq c\|x - y\|^2$. Using the strong elliptical property of L as above,

$$
\begin{aligned}
E(V_j'(x) - V_j'(y))^2 &\leq c \sum_{k \in Z^4} \alpha_k^2 \|2\pi k\|^2 \frac{|e^{j2\pi\langle k,x\rangle} - e^{j2\pi\langle k,y\rangle}|^2}{\|k\|^{4q}} \\
&\leq c\|x - y\|^2 (c_1 + \sum_{k \in Z^4/F} \frac{\|k\|^{2(p+2)}}{\|k\|^{4q}}) \ .
\end{aligned}
\tag{96}
$$

The series converges for $4q - 4 - 2p > 4$. ∎

6.2 Large deformation templates

Large deformation template construction proceeds assuming $(\Omega, \mathcal{I}, \mathcal{H})$ is a homogeneous anatomy, and we assume the background space $\Omega \doteq [0,1]^3$ is compact, corresponding to a closed volume. Given are a set of $N+1$ landmarked brains or imaged brains $I_n, n = 0, \ldots, N \subset \mathcal{I}$ defined on Ω. The goal is to construct a brain $I_{temp} \in \mathcal{I}$ which most closely represents the population.

Empirical estimation of the templates for the various subpopulations of interest use the ideas of *minimum mean-squared estimation*(MMSE) for generating the template. A crucial aspect of the approach is that the template coordinate system is generated from averages of transformations, not from averages of the image data itself. The rational behind this is that averages of images have no convincingly biological interpretation in contrast to averages of deformations that generate another deformation.

For large deformation maps, empirical averages of the vector fields will not generate diffeomorphisms. This behooves us to introduce a large deformation setting for understanding the maps. This is most naturally accomplished in the vector field setting for understanding the velocity fields.

Let $\{v_n, n = 1, \ldots, N\}$ be the set of velocity fields mapping I_0 to $I_n, n = 1, \ldots, N$. Then the *large deformation template* is defined to be $\bar{h}I_0$ where $\bar{h}$ satisfies the ODE with $\bar{v} = \sum_{i=1}^N v_i$:

$$
\frac{\partial \bar{h}(x,t)}{\partial t} = -\nabla_x^t \bar{h}^t(x,t) \bar{v}(x,t) \ , \ h(x,0) = x \ , \ t \in [0, T] \ .
\tag{97}
$$

If $v(x,t)$ is continuously differentiable vector field on $[0,1]$ then by the existence and uniqueness theorem of O.D.Es, the solution exists and is uniquely determined by the velocity field $v(x,t)$ and the initial condition $h(0,x)$ furthermore it defines a unique diffeomorphism of Ω via the solution to the above O.D.E. By a continuously differentiable vector field, following [3] we mean that each of the coordinate functions $v_i(x,t), i = 1, 2, 3$ are continuously differentiable with respect to $x_i, i = 1, 2, 3$ and t.

Joshi [16] has constructed a large deformation average under the landmark diffeomorphisms. Examine the whole macaque cryosection brains $87A, 90C, 93G$ shown in Fig. 12 in which the gyri and associated sulci have been labeled. The sulci and gyri can be defined precisely in terms of the geometrical properties of the cortical surface using the notions of ridge curves and crest lines (extremal points of curvature) following [70, 24]. In Fig. 12 each of the sulci have been identified and placed into the whole brains discretized to 16 points.

Following Joshi [16] the brain $87A$ was selected as the provisory template I_0 and was mapped on to the two brains $90C$ and $93G$, using the large deformation landmark matching algorithm yielding the velocity fields which when integrated through the ODE generate the map of $I_0 \doteq 87A$ onto the two brains $90C, 93G$. An average velocity field $\bar{v}$ was constructed by averaging over the velocity field maps, and the O.D.E defined in Eqn. 5 was solved yielding the average diffeomorphic transformation $\bar{h}(x, T)$. Panel 4 of Figure 12 is the resulting the large deformation template representing the population. Notice that as the velocity fields used to generate the empirical mean of the population were driven only by the identified landmarks on the fundus of the fundus beds in the sulci in the three brains the external Calcarine fundus curve present only in 90C and not used for generating the transformations is not represented in the large deformation template shown in Figure 12.

Shown in Figure 13 are six sections through the large deformation template generated from the population.

7 Surfaces

Examine two-dimensional smooth surface manifolds $\mathcal{S}$ such as the cortical and hippocampul surface.
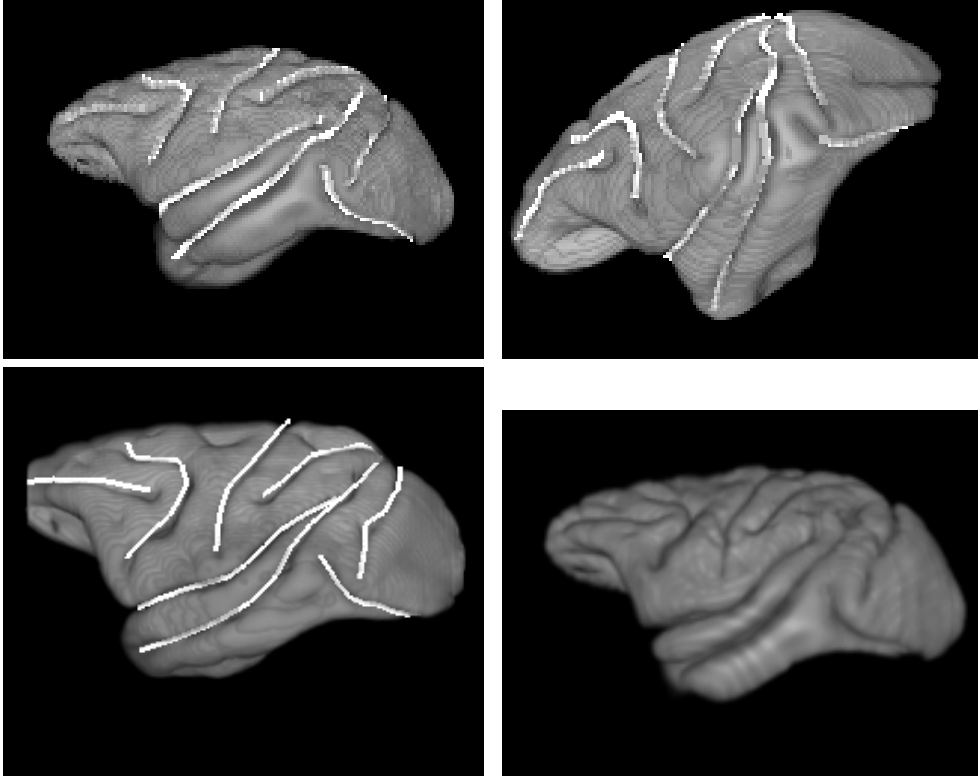
Figure 12: Panels 1,2,3 shows three brains 87A , 90C and 93G in which the gyri and associated sulci which have been labeled. The sulci and gyri are defined precisely in terms of the geometrical properties of the cortical surface using the notions of ridge curves and crest lines (extremal points of curvature). Panel 4 shows the average template brain.

7.1 Large Deformation Mapping of Surfaces

The goal is to define a distance between elements in the orbit. Define the space of anatomical imagery to be two dimensional manifold surfaces under the group of diffeomorphisms restricted to actions on the template surface:

$$\mathcal{S} \doteq \mathcal{H}S_{temp} \doteq \{hS_{temp}, h \in \mathcal{H}\} \text{ where } hS_{temp} \doteq \{h(x) : x \in S_{temp}\} . \tag{98}$$

Notice, the action of $\mathcal{H}$ on S_{temp} is restricted to the surface. This is assumed throughout. The surfaces are triangulated graphs with attached tangent spaces and shape operator associated with locally quadratic charts, as described in with the geometry of the quadratic charts transforming in the standard way with the Jacobian of the mapping transforming the tangent space, and curvature transformed by the Hessian of the transformation [20]. This forms the basis for the distance between the family of surfaces.

Assume S to be a smooth two-dimensional differentiable sub-manifold of class two. Then each point $x \in S$ has a tangent space and orthonormal basis of the tangent plane $T_{x_i}(S)$ and $b_{x_i}^{(1)}, b_{x_i}^{(2)}$ at each point x_i, with the unit normal n_{x_i}. The quadratic form C_{x_i} and tangent space basis b_1, b_2 are estimated at each point on the surface by a minimum mean squared error algorithm described in section 3.4.2.

Once the local charts are established at each vertex, intrinsic geometrical properties of the surface can be studied. In defining distances between surface we exploit the principal curvatures κ_1, κ_2 of C_{x_i} describing the maximum and minimum curving degrees of the surface. corresponding to the eigen values of the 2×2 shape operator matrix C_{x_i}. Let S, S' be the template and target surfaces with a map restricted to the surface taking the point on the template $x \in S$ to a point on the target surface $x' \in S'$:

$$h \in \mathcal{H} : x \in S \mapsto h(x) = x' \in S' .$$

Let the tangent planes at the points be $T(x)$ and $T(x')$. How should the distance between surfaces be defined?

We want the map h to carry the differential-geometric features modulo the special euclidean group $\mathbf{SE}(3)$. The latter means that we are operating in equivalences relative to the Euclidean group since a local translation/rotation should
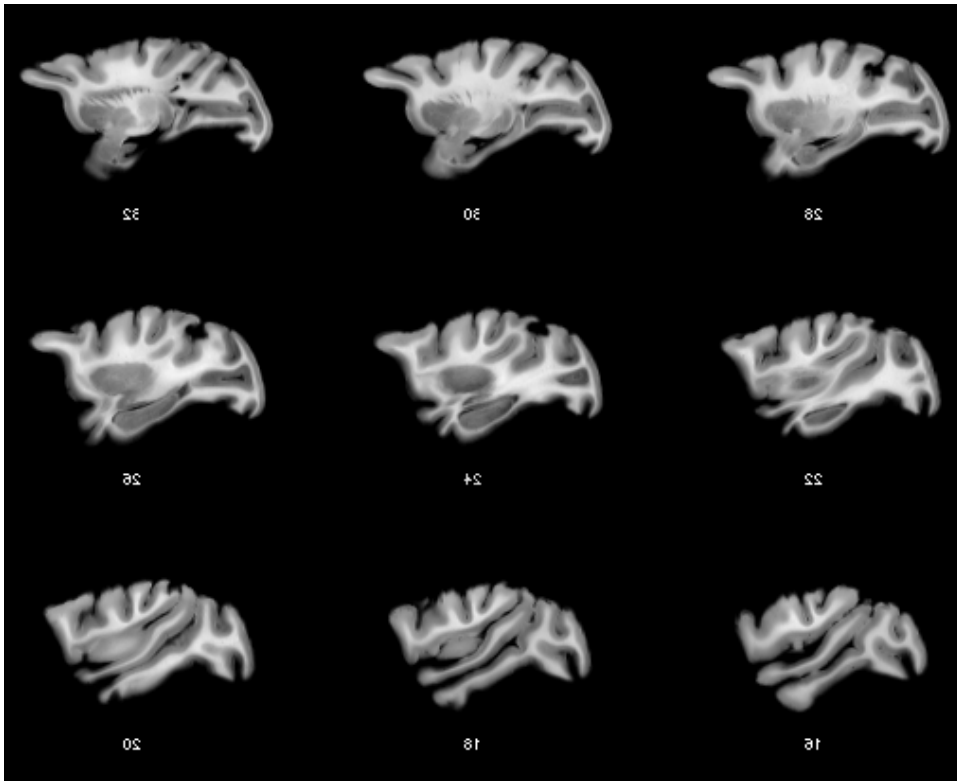
Figure 13: The figure shows sections from the large deformation template brain generated from the population 87A,90C and 93G by composing the diffeomorphism generated from the average of velocity fields with the initial anatomy $\bar{h}I_0$.

not affect the value of the criterion. Thus we are constructing a representation invariant to this group. The translation, with three degrees of freedom, should take $x \mapsto x'$, followed by a rotation, with two degrees of freedom, that takes $T_x \rightarrow T_{x'}$. Since the dimension $\dim \mathbf{SE}(3) = 6$ there remains one degree of freedom for the rotation around the normal to the tangent element. The infinitesimal properties of order two of the surface element after this transformation can be expressed through two homogeneous quadratic forms, say with matrices M and M' representing curvature. Use the Hilbert-Schmidt norm between M and the rotated (around the normal) M', asking that it be minimized, the Hilbert-Scmidt norm of a symmetric matrix $M = (m_{ij})$ becomes $\|M\|_{HS} = \sqrt{\sum_{i,j} m_{ij}^2} = \sqrt{tr(M^2)}$. We should now find the rotation, say $O \in \mathbf{SO}(2)$, that makes the infinitesimal structures as close as possible. The following solves this extremum problem, providing the proper energy for the surface matching.

Lemma 1 *With the symmetric 2×2 matrices M and M' having the principal curvatures the eigenvalues $\kappa_1 \geq \kappa_2$ and $\kappa_1' \geq \kappa_2'$, then*

$$\min_{O \in \mathbf{SO}(2)} \|M - OM'O^t\|_{HS}^2 = \sum_{k=1}^{2} (\kappa_k - \kappa_k')^2 \ .$$

Proof: Assume without loss of generality that M is in diagonal form $M = diag[\kappa_1, \kappa_2]$, then using the property that the trace is invariant with respect to rotations gives

$$
\begin{aligned}
\min_{O \in \mathbf{SO}(2)} \|M - OM'O^t\|_{HS}^2 &= \min_{O \in \mathbf{SO}(2)} tr[(M - OM'O^t)^2] \\
&\overset{(a)}{=} \min_{O \in \mathbf{SO}(2)} tr(M^2) + tr((M')^2) - 2tr[MOM'O^t] \quad (99) \\
&= \max_{O \in \mathbf{SO}(2)} tr[MOM'O^t] \quad (100)
\end{aligned}
$$

and (a) follows from the invariance the trace with respect to rotations. The last trace equals $\sum_{k=1}^{2} \kappa_k x_k$ where $x_k = (OM_2O^t)_{kk}$. But whatever O is we have

$$
\begin{aligned}
\sum_{k=1}^{2} x_k &= tr(M') = \sum_{k=1}^{2} \kappa_k' = constant \ , \quad (101) \\
&= \kappa_1 x_1 + \kappa_2 (constant - x_1) = (\kappa_1 - \kappa_2) x_1 + \kappa_2 \times constant \ . \quad (102)
\end{aligned}
$$

This implies that to maximize x_1 should be chosen as large as possible, since

$$x_1 = uM'u^t; u = O \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ ,$$

with x_1 which cannot be made larger than κ_1'. Then $x_2 = \kappa_2'$ and the stated result follows by inserting these values into Eqn. 100. ∎

This forms the basis for our surface matching. Define $\| \cdot \|^2$ to be the norm-squared on functions indexed over the surface S.

Problem Statement 2 (Surface Matching) *Given are surfaces S, S' with curvature fields $\kappa_k(\cdot), \kappa_k'(\cdot), k = 1, 2$. The optimal match $\hat{h}$ of S to observation S' becomes*

$$
\begin{aligned}
\hat{h}(x, T) &= \int_0^T -\nabla_x^t \hat{h}(x, t) \hat{v}(x, t) dt + x \text{ where} \quad (103) \\
\hat{v}(\cdot) &= \arg \min_{v \in \mathcal{V}} \|v\|_{\mathcal{V}}^2 + \int_S \sum_{k=1}^{2} (\kappa_k(h(x)) - \kappa_k'(x)) dm(x) \ , \quad (104)
\end{aligned}
$$

where m is the surface measure for S. For landmark matching, replace the curvature penalty by landmark distances.

As described below, we have implemented Eqn. 104 with landmarks only, not with curvature maps.

7.2 Cortical Topography and Flat Maps

The mammalian cerebral cortex has the form of a layered, highly convoluted thin shell of gray matter surrounding white matter, and is one of the most striking features of the brain. The cortex contains a complex mosaic of of anatomically and functionally distinct areas which play a tremendously important role in understanding brain functions[78]. To aid in the understanding of the geometry and the functional organization of the cortex, Van Essen [65, 38] has championed the approach of visualizing and comparing brains via cortical flat maps. Figure 14 shows the cortical surface and the associated flat map generated by Van Essen, depicting the complex functional topography of the cortex. The left panel shows a colored cortical surface representing the various areas of the cortex; the middle panel shows the Van Essen group's flattened representation depicting the areas in the Felleman-Van Essen map.

Figure 14: Left panel shows the whole cortical surface labelled in Felleman Van Essen 79O brain. Middle panel shows the flattened 79O brain with cortical areas. Right panel shows the Brodman partitioning. Surfaces and data are courtesy of Van Essen et al. 1997.

To accommodate individual variation in the cortical topography across brains, we use landmark transformation tools as developed in Joshi [16] applied to the flat maps of various individual cortical surfaces [23]. Shown in the right panel of Figure 14 is the now classic Brodman areas displayed in a flattened representation. Shown in the top row of Figure 15 is an illustration of the geometrical features associated with the curvature of the cortical surface such as the sulci and the fundi that were identified on both the flat maps. These features are used as the landmarks for driving the deformation algorithms. Shown in Figure 16 is the result of the deformation process. The left panel of Figure 16 shows the deformed flat map corresponding to the partitioning schemes by Brodmann. Shown in the right panel is the overlay of the deformed partitioning schemes by Brodmann on the partitioning scheme by Felleman and Van Essen. The bottom row of figure 16 shows the correspondence of the geometrical fundus features after transformation with fixed variance in the landmark matching. Right panel shows the resulting transformation of the grid in the large deformation surface matching.

8 Sulcus and Gyral Curves of the Neocortex

Modern whole brain cryosection imaging provides excellent high resolution data required for study of such anatomical features of cortex [133, 78] such as the arrangement of the sulcal fissures visible throughout the cortical surface of a mammalian with major sulci and gyri now being catalogued in human atlases which and becoming available [38, 23]. Computational metrics defined by cortical geometry such as geodesic length is drawing the attention of the neuroscience community in terms of the role of wiring length in the general layout of the nervous system [66]. Despite their anatomic and functional significance, the gyri, sulci and many stable cortical pathways consistently appearing in all normal anatomies exhibit pronounced variability in size and configuration [134]. Methods are beginning to appear for characterizing their variation [135]. We now examine methods for defining extremal curves on triangulated graphs via dynamic programming, constructing probability laws for characterizing their shape and variability, and matching curves across brains.

8.1 Dynamic Programming Generation of Sulcus, Gyral and Geodesic Curves

The sulci and gyri exhibit strong features associated with their extrema of bending. The deepest beds of the sulci are called the fundus beds; associated with the gyri are the crowns. The automated generation of fundus beds along with
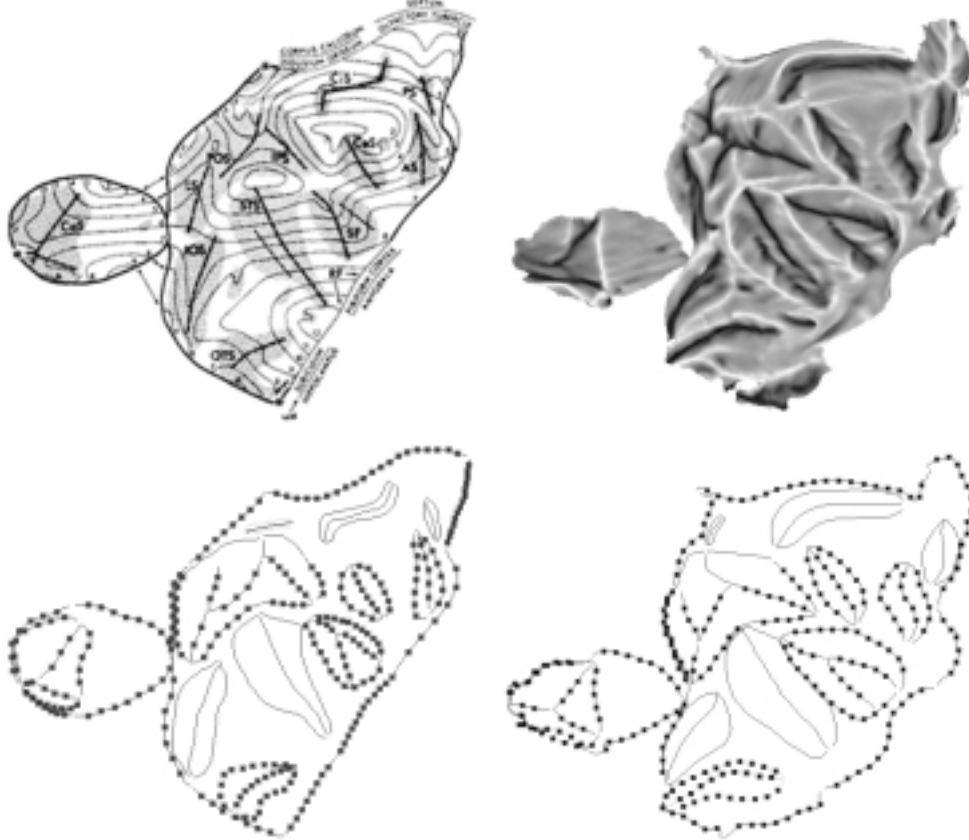
Figure 15: Top row: Figure shows the cortical geography as indicated by the fundi in Broadmann's maps (left panel) and the mean curvature field for brain 79O Canonical Van Essen brain (right panel). Bottom row: The geometrical fundus bed features identified on the flat maps and used as landmarks for the deformation process. Taken from Van-Essen et al. 1997.

geodesics will now be discused, with the curvature sign reversed to generate the gyral crowns.

Empirical evidence [135, 38, 23] suggests that in most part of their length fundus beds resemble crest lines (see [67, 136]) corresponding to points where the maximal absolute principal curvature has a local maximum. There exist algorithms for extracting ridge and crest lines from surface geometries [67, 69, 68], from which zero tracing algorithms involving computing gradients of the curvature function are constructed. Crest lines and curves [67, 136, 68] are the loci of points $x \in S$ where the maximal absolute principal curvature $\kappa_{max}(x)$ has a maximum. Let the surface have local quadratic approximation $x(u,v)$ locally around $x \in S$ from Eqn. 36, then

$$x(u,v) = x_i + u b_{x_i}^{(1)} + v b_{x_i}^{(2)} + ((u,v)C_x(u,v)^t)n_{x_i} \; , \tag{105}$$

with eigenvalues and eigenvectors

$$C_x t_i \;\; = \;\; \kappa_i t_i \; , \; i = 1, 2 \; ; \tag{106}$$

$$\kappa_{\arg\max_{i=1,2} |\kappa_i(x)|} \; , \; t_{\arg\max_{i=1,2} |\kappa_i(x)|} \; . \tag{107}$$

At these points, $\nabla \kappa_{max} \cdot t_{max} = 0$ for t_{max} the principal unit direction corresponding to the maximal principal curvature κ_{max}. Such zero tracing methods are sensitive to noise. We take an alternate approach based on dynamic programming for tracking optimal trajectories on surfaces similar to that done in boundary and artery tracking [137, 138]. The problem of tracking such curves is posed as a control/optimization problem. We search for curves that pass through regions of highest maximal curvature joinng the prespecified start and end points (s, t) in the surface. For noise immunity, instead of finding the extremum of principal curvature using higher derivatives of curvature, we define a sequentially additive energy associated with candidate curves and use dynamic programming for its minimization. To apply dynamic programming, we define the cost of a candidate curve $\alpha(s, t) \in P(s, t)$ to be $\int_{\alpha(s,t)} (\kappa_{max}(x) - \mathcal{K})^2 dx$, with $\mathcal{K}$ assigned
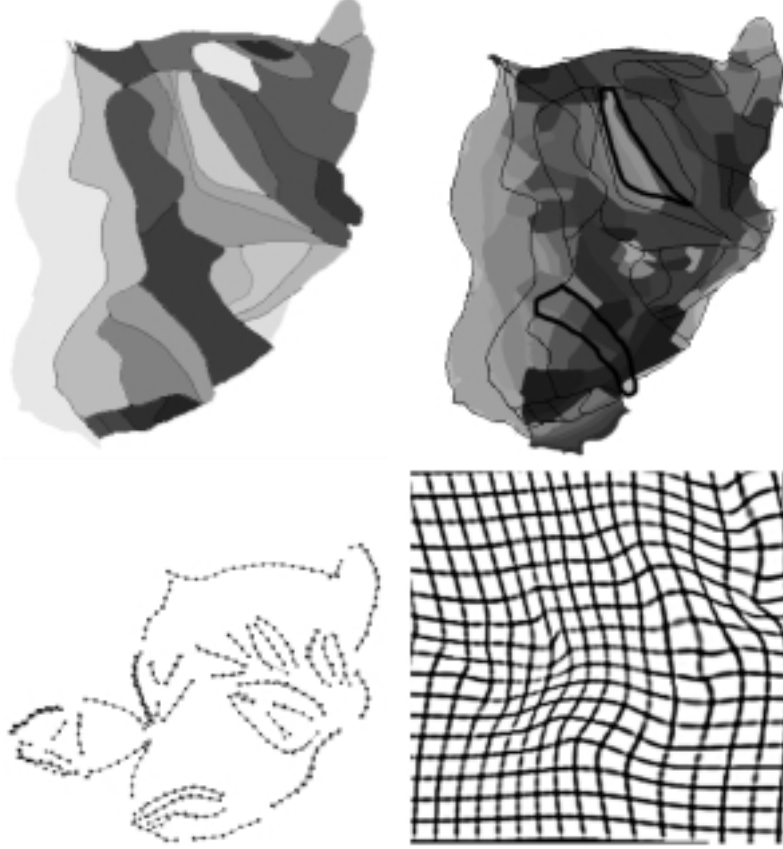
Figure 16: Top row: Left panel shows Brodmann map after deformation to 79O Felleman and Van Essen brain; right panel shows overlay of Brodmann on the partitioning scheme by Fellman and Van Essen. Bottom row: Left panel shows the geometrical fundus features after transformation with fixed variance in the landmark matching. Right panel shows the resulting transformation of the grid in the large deformation surface matching. Taken from Van-Essen et al. 1997.

the largest maximal curvature on the surface, and minimize over all such paths on the triangulated graph representation of the surface. We note, for a surface symmetrical about a crest line, i.e. one where t_{max} is perpendicular to the crest line, minimizing the above function gives $(\kappa_{max}(x) - \mathcal{K})\nabla\kappa_{max} \cdot t_{max} = 0$ implying $\nabla\kappa_{max} \cdot t_{max} = 0$ which is precisely the equation for the crest line. For regions of the sulcus where the basin is flat, κ_{max} is constant, and the minimizer of the above functional produces shortest paths through these regions.

For generating *geodesics*, we adapt the continuum definition associated with length on the surface as measured by the integral of the norm of the tangent vector along the curve. If s and t are points of a smooth connected surface $S \subset \mathbb{R}^3$, the intrinsic distance (Riemannian length) $\rho(s, t)$ from s to t in S is the lower bound of the lengths of these curves. The curve $\hat{\alpha}$ for which the minimum length is acheived is called a geodesic. Throughout we denote the triangulation of points on the surface as i, j, and the set of coordinates in $\mathbb{R}^3$ taken by the sites of the graph and or positions of the triangle vertices as $x_i, x_j \in \mathbb{R}^3$.

Definition 3 *Given a two-dimensional triangulation of the surface $\{x_i \in S\}$ define the **platelet** $\mathcal{P}_i$ of point i as the set of triangles (with index-triples (j_1, j_2, j_3) specifying their vertices) sharing x_i as a common vertex $\mathcal{P}_i = \cup\{(j_1, j_2, j_3)|x_i = x_{j_1}$ or $x_i = x_{j_2}$ or $x_i = x_{j_3}\}$.*

Define a path on the surface $\alpha(s, t)$ routed and terminated respectively in nodes s, t on the surface as

$$\alpha(s, t) \doteq (s = j_1, j_2), (j_2, j_3) \dots, (j_{k-1}, j_k), \dots, (j_{N-1}, t = j_N) \text{ , such that } j_k \in \mathcal{P}_{j_{k-1}}, \forall k \text{ ,}$$

and the collection of all paths connecting (s, t) as $\alpha(s, t) \in P_{s,t}(S)$.

Define the N-length discrete geodesic and discrete fundus bed as cost minimizing paths given by

$$\text{fundus}(s,t) \;\doteq\; \arg \min_{\alpha(s,t)\in P_{s,t}(S)} \sum_{k=1}^{N} d(j_k, j_{k+1}) \tag{108}$$

$$\text{where } d(j_k, j_{k+1}) \;\doteq\; \left(\frac{(\kappa_{max}(x_{j_k}) - \mathcal{K})^2 + (\kappa_{max}(x_{j_{k+1}}) - \mathcal{K})^2}{2} + \frac{(\kappa_{max}(x_{j_k}) - \kappa_{max}(x_{j_{k+1}}))^2}{6} \right) \|x_{j_k} - x_{j_{k+1}}\| \;;$$

$$\text{geodesic}(s,t) \;\doteq\; \arg \min_{\alpha(s,t)\in P_{s,t}(S)} \sum_{k=1}^{N} d(j_k, j_{k+1}) \tag{109}$$

$$\text{where } d(j_k, j_{k+1}) \;\doteq\; \sqrt{(x_{1,j_k} - x_{1,j_{k+1}})^2 + (x_{2,j_k} - x_{2,j_{k+1}})^2 + (x_{3,j_k} - x_{3,j_{k+1}})^2} \;.$$

Dynamic programming. For generating these curves and geodesics we follow Khaneja [70, 24] using dynamic programming adapted to optimization on triangulated surfaces. Denote the finite state space S of size $\|S\| = N$; on these triangulated graphs the positions of the nodes of the surface itself are the states. The goal is to compute optimal shortest paths between the specified initial states s and the final state t. Assuming that the optimal path has no more that K nodes, the total number of paths of length K between points s and t are of the order N^K. If the cost is additive over the length of the path, dynamic programming reduces the complexity of the search algorithm to order of KN^2. Let $c^k(x_k, x_{k+1})$ denote the cost incurred for the transition from state $x_k \in S$ to $x_{k+1} \in S$ at each time k. Suppression of k dependence in $c(i,j)$ means the cost is independent of time. We shall assume that $c(i,j) \geq 0$, and arcs of infinite cost $c(i,j) = \infty$ signify that there is no arc from node i to node j. An optimal path need not have more than N arcs (number of nodes in the graph) and hence take no more than N moves. We formulate the problem as one of finding the optimal path in exactly N moves allowing degenerate moves from a node i to itself with cost $c(i,i) = 0$. The degenerate moves signify that the length of the path may be less than N.

The efficiency of dynamic programming on the triangulated graphs representing the surface is that the states spaces $S_k \subset S$ can be dynamically defined and of reduced complexity. Curves passing through a point on the graph must pass through one of it neighbors (analogous to being in the tangent space for the continuum representation).

Algorithm 5 (Dynamic Programming Algorithm) *Denote the optimal cost for getting from node i to node t in $(N - k)$ moves as $J_k(i), i \in S$, $k = 0, 1, \cdots, N - 1$. Then the optimal N-length path $J_0(i)$ from i to t is given by the final step of the following algorithm,*

$$\text{with } J_{N-1}(i) \;=\; c^{N-1}(i,t) \;,\text{ and} \tag{110}$$

$$J_k(i) \;=\; \min_{j=1,\cdots,N} \{c^k(i,j) + J_{k+1}(j)\}, \;\; k = 0, 1, \cdots, N - 2 \;, i \in S \;. \tag{111}$$

Define the state spaces dynamically $S_{N-1} = \{i | i \in \mathcal{P}_t\}$, $S_k = \{i | i \in \mathcal{P}_j, j \in S_{k+1}\}$, and implement the algorithm according to Initialize: $J_k(i) \leftarrow \infty \quad i \neq t$, for all k, $S_N \leftarrow t$, $J_k(t) \leftarrow 0$;
For $k \leftarrow N - 1$ down to 0 do

$$S_k \leftarrow \{i \mid i \in \mathcal{P}_j, j \in S_{k+1}\} \;,\text{set } c^k(i,j), j \in S_{k+1}, i \in S_k \;, \tag{112}$$

with

$$J_{N-1}(i) \;=\; c^{N-1}(i,t), i \in S_{N-1} \;, \tag{113}$$

$$J_k(i) \;=\; \min_{j\in\{S_{k+1}\bigcap \mathcal{P}_i\}} \{c^k(i,j) + J_{k+1}(j)\}, \;\; i \in S_k, \quad k = 0, 1, \ldots, N - 2 \;. \tag{114}$$

Theorem 8 *1. **Geodesic generation.** Given the costs for transition*

$$c^k(i,j) \;=\; d(i,j) \doteq \sqrt{(x_{1,i} - x_{1,j})^2 + (x_{2,i} - x_{2,j})^2 + (x_{3,i} - x_{3,j})^2} \;, \; j \in \mathcal{P}_i \tag{115}$$

$$c^k(i,j) \;=\; \infty \text{ for } j \notin \mathcal{P}_i \;,$$

then the length $J_0(s)$ defined in the algorithm corresponds to a geodesic shortest path (not necessarily unique) between nodes s and t:

$$\text{geodesic}(s,t) \doteq \arg \min_{\alpha(s,t)\in P_{s,t}(S)} \sum_{k} \sqrt{(x_{1,k} - x_{1,k+1})^2 + (x_{2,k} - x_{2,k+1})^2 + (x_{3,k} - x_{3,k+1})^2} \;, \quad .$$

2. Fundus curve generation. *Given the costs for transition*

$$c^k(i,j) = \left(\frac{(\kappa_{max}(x_i) - \mathcal{K})^2 + (\kappa_{max}(x_j) - \mathcal{K})^2}{2} + \frac{(\kappa_{max}(x_i) - \kappa_{max}(x_j)^2}{6} \right) \|x_i - x_j\| , \quad j \in \mathcal{P}_i , \quad (116)$$

$$c^k(i,j) = \infty , \quad j \notin \mathcal{P}_i,$$

then the length $J_0(s)$ defined in the algorithm corresponds to an optimal fundus bed path between nodes s and t:

$$\hat{\alpha}(s,t) \doteq \arg \min_{\alpha(s,t) \in P_{s,t}(S)} \sum_k \left(\frac{(\kappa_{max}(x_k) - \mathcal{K})^2 + (\kappa_{max}(x_{k+1}) - \mathcal{K})^2}{2} + \frac{(\kappa_{max}(x_k) - \kappa_{max}(x_{k+1}))^2}{6} \right) \|x_k - x_{k+1}\| .$$

For a proof see [70, 24] which follows since the costs are additive over the curves.

Notice the cost of a discrete curve $\alpha(s,t)$ connecting (s,t) is defined using the corrected trapezoid rule approximation for the curvature integral $\int_{\alpha(s,t)} (\kappa_{max}(x) - \mathcal{K})^2 dx$.
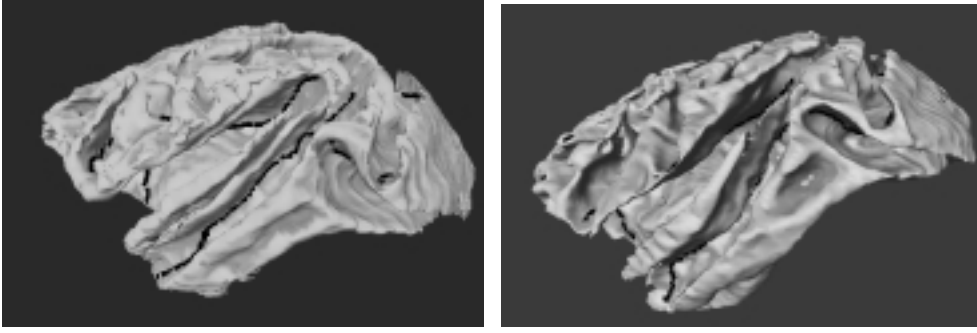

Figure 17: Macaque monkey fundus curves generated via dynamic programming and costs of Eqn. 116. Top row shows Superior Temporal and Sylvian fissure fundus curves generated in brains 87A and 93G. Bottom row shows the fundus curves of the Central sulcus, Arcuate sulcus, Principal sulcus and Lunate sulcus generated on brain 95CL. Curves generated as extremal points of principal curvature via dynamic programming.

The dynamic programming algorithm was used to extract fundus curves from the surface of the macaque monkey brains $87A, 95CL$ from David Van Essen's laboratory. Shown in the top row of figure 17 are the Superior Temporal and Sylvian fissure fundus curves generated in brains 87A and 93G generated using the costs of Eqn. 116. Bottom row show the fundus curves of the Central sulcus, Arcuate sulcus, Principal sulcus and Lunate sulcus generated on 95CL. Curves generated as extremal points of principal curvature via dynamic programming.

Shown in Figure 18 are results from the Visible Human cortex generated in David Van Essen's laboratory. Panel 1 shows the dynamic programming extracted superior temporal fundus on the Visible Human cortex generated by Van Essen and Drury [23]. Panel 2 shows choosing multiple start and end points for the dynamic programming solution finding common flows on the surface. The paths are illustrated on flattened Drury-Van Essen Visible Human flat map illustrating robustness of the solution. The algorithm was run for different choices of the starting and end points. Notice how the trajectories initiated at different starting points merge into a common trajectory, emphasizing the optimality of the path. The endpoint control formulation also allows for dealing with the presence of bifurcation, allowing for tracking of individual branches of fundus beds. Panel 3 shows the dynamic programming generation of the superior temporal
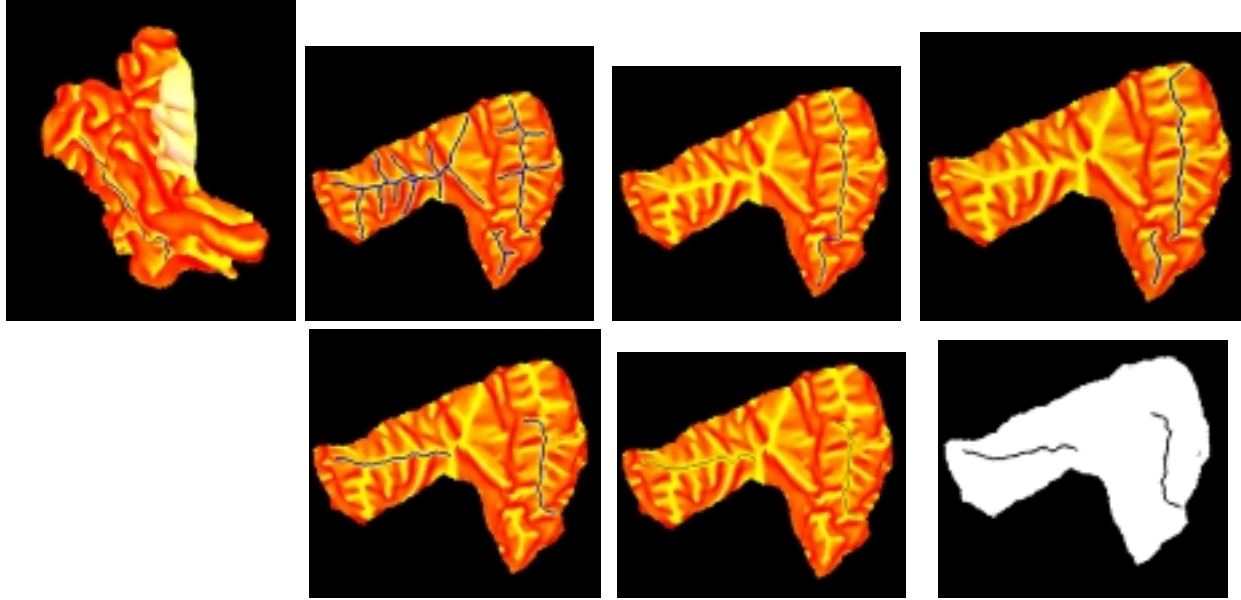
Figure 18: Visible human fundus curves generated via dynamic programming and costs of Eqn. 116. Top row: Panel 1 shows the dynamic programming extracted fundus on the Visible Human cortex; panel 2 shows choosing multiple start and end points for the dynamic programming solution finding common flows on the flattened Drury-Van Essen Visible Human flat map illustrating robustness of the solution; panel 3 shows the dynamic programming generation of the superior temporal sulcus jumping across the break connecting the start and end points which were manually selected. panel 4 shows the dynamic programming solution choosing two start and end points giving a break in the superior temporal sulcus; Bottom row: Panel 6 shows the main brainches of the Superior Temporal and Sylvian Fissure fundus beds extracted using dynamic programming rendered on the Drury-Van Essen Visible Human flat map; panel 7 shows the main branches of hand drawn Superior Temporal and Sylvian Fissure fundus beds drawn by Van Essen; panel 8 shows an overlay of the hand and automated fundus curves.

sulcus jumping across the break connecting the start and end points which were manually selected. These are rendered on the Visible Human cortical flattened surface. The bottom row panel 4 shows the dynamic programming generation of the superior temporal sulcus when two start and end points are chosen allowing for a break.

The bottom row of Figure 18 shows comparison of automated generation of fundus curves and curves generated by David Van Essen by hand. Panel 6 shows the main branches of the Superior Temporal and Sylvian Fissure fundus curves extracted using dynamic programming rendered on on a Drury-Van Essen flat map of the Visible Human [23]. Panel 7 shows the branches of the Superior Temporal and Sylvian Fissure fundus beds hand generated by Van Essen on the flat map. For comparison of accuracy, panel 8 (bottom row) shows an overlay of the hand and automatic generation of the fundus curves illustrated on the flat map [23]. Notice the exact correspondence.

Shown in Figure 19 are results from generating automatically the fundus curves of the Superior Temporal Gyrus (STG) region from the Visible Male. The cortical surface was generated using the semi-automated surface generation utilities of M. Joshi [139]. Left panel shows the fundus curve of the Superior Temporal Sulcus (STS), middle panel the fundus curve of the Sylvian Fissure (SF) and right panel shows the crown of the STG. The curves were generated by choosing the start and end points for the dynamic programming solution.

Shown in the left panel of Figure 20 are eight geodesics constructed on the cortical surface of the macaque after manual selection of start and termination points. The gray matter defining the neocortical surface was hand segmented, from which a triangulation was generated using Marching cubes [122] to have approximately 15000 nodes. The middle panel shows 6 points which corresponds to origins of the various fundus beds which were used to compute geoedesic distance: 1=superior point of Arcuate Sulcus, 2=superior point of Central Sulcus, 3=inferior point of Central Sulcus, 4=anterior point of Intra Parietal Sulcus, 5=inferior point of Arcuate Sulcus and 6=posterior point of Principal Sulcus. The right panel shows the shortest geodesic distances that were computed in millimeters on the cortical surface. For example, the geodesic distance from the Arcuate Sulcus to the 5 other points in millimeters were $d(1,2) = 47.07mm$, $d(1,3) = 63.62mm$, $d(1,4) = 68.75mm$, $d(1,5) = 39.29mm$, $d(1,6) = 22.39mm$.
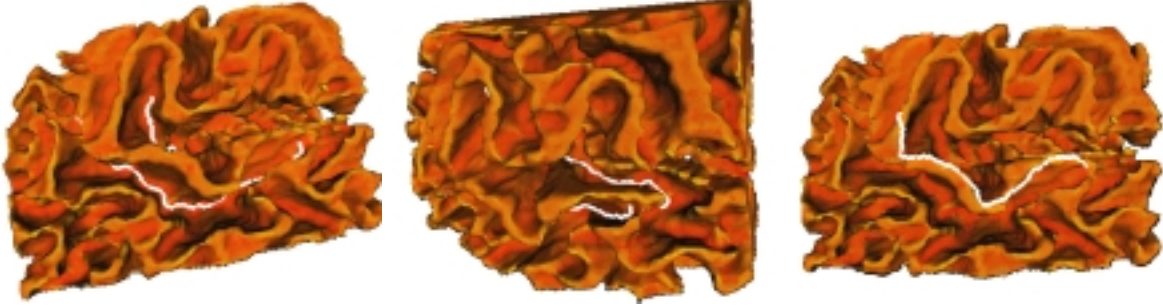
Figure 19: Figure showing the Superior Temporal Gyrus (STG) region in the Visible Male. Left panel shows the fundus curve of the Superior Temporal Sulcus (STS), middle panel the fundus curve of the Sylvian Fissure (SF) and right panel shows the crown of the STG.

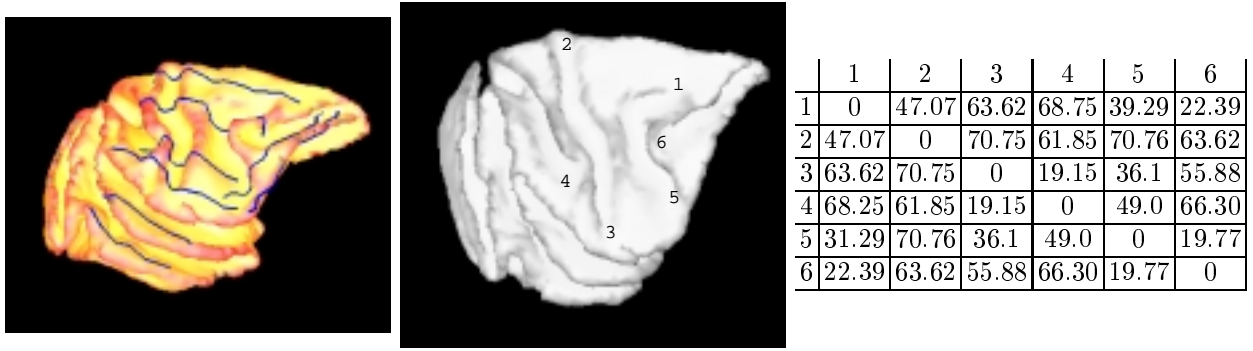|	1	2	3	4	5	6
1	0	47.07	63.62	68.75	39.29	22.39
2	47.07	0	70.75	61.85	70.76	63.62
3	63.62	70.75	0	19.15	36.1	55.88
4	68.25	61.85	19.15	0	49.0	66.30
5	31.29	70.76	36.1	49.0	0	19.77
6	22.39	63.62	55.88	66.30	19.77	0

Figure 20: Panel 1 shows eight geodesics generated on the neocortex by picking the start and end points manually. Panel 2 depicts geographical landmarks on the macaque cortex; labels 1, 2, 3, 4, 5, 6. Panel 3 shows a table of Riemannian distances in millimeters between the predefined points. Data taken from the laboratory of David Van Essen, Washington University.

8.2   Curve Matching via Frenet Induced Distance Metrics

We induce distance measures between curves following Khaneja [140] using the Frenet representation [136, 141], of curves in terms of velocity, speed, and torsion. The Frenet distance measures provide the basis for diffeomorphic matching. This is in the same spirit as the beautiful work on diffeomorphic matching of curves by Younes [71]. Let $x(\cdot)$ be a smooth curve, $x(s), s \in [0,1]$ parameterized via arc-length, and associate the orthogonal frame corresponding to the tangent vector field, normal field, and binormal $O(s) \doteq (T(s), N(s), B(s)), s \in [0,1]$ satisfying the ODE via the orthogonal frame representation $O(s) = [T(s), N(s), B(s)]$, where $\frac{dO(s)}{ds} = F(s)O(s), O(0) = I$, and defined via the Frenet velocity, curvature, torsion representation:

$$\underbrace{\begin{pmatrix} \frac{dT(s)}{ds} \\ \frac{dN(s)}{ds} \\ \frac{dB(s)}{ds} \end{pmatrix}}_{\frac{dO(s)}{ds}} = \underbrace{\begin{pmatrix} 0 & \alpha(s)\kappa(s) & 0 \\ -\alpha(s)\kappa(s) & 0 & \alpha(s)\tau(s) \\ 0 & -\alpha(s)\tau(s) & 0 \end{pmatrix}}_{F(s)} \underbrace{\begin{pmatrix} T(s) \\ N(s) \\ B(s) \end{pmatrix}}_{O(s)} . \tag{117}$$

Here $F(\cdot)$ is the skew-symmetric matrix describing the flow of the orthogonal frame of the curve through it's tangent space, and the speed, curvature and torsion functions are given by

$$\alpha(s) = \|\dot{x}(s)\| \, , \; \kappa(s) = \frac{\|\dot{x}(s) \times \ddot{x}(s)\|}{\|\dot{x}(s)\|^3} \, , \; \tau(s) = \frac{(\dot{x}(s) \times \ddot{x}(s)) \cdot \dddot{x}(s)}{\|\dot{x}(s) \times \ddot{x}(s)\|} \, ; \tag{118}$$

for unit speed curves it is given by

$$\alpha \;=\; 1 \;,\; \kappa(s) = \|\frac{d^2 x(s)}{ds^2}\| \;,\; \tau(s) = \frac{1}{k^2(s)}\det\begin{pmatrix} \frac{dx_1(s)}{ds} & \frac{dx_2(s)}{ds} & \frac{dx_3(s)}{ds} \\ \frac{d^2 x_1(s)}{ds^2} & \frac{d^2 x_2(s)}{ds^2} & \frac{d^2 x_3(s)}{ds^2} \\ \frac{d^3 x_1(s)}{ds^3} & \frac{d^3 x_2(s)}{ds^3} & \frac{d^3 x_3(s)}{ds^3} \end{pmatrix} . \tag{119}$$

Identify the curves with their velocity field $F(s), s \in [0,1]$. Define the metric distance between curves x, x' as $\rho(x, x')$ by identifying them with their flow through the tangent space. Consider two simple regular curves $x(s), x'(s), s \in [0,1]$, then we are looking to establish a distance function $\rho(x, x')$, x, x' smooth regular curves, that is invariant to spatial position and orientation in the space. Define the diffeomorphisms restricted to the curve as $\phi \in \Phi$ on the regular curves $x(\cdot)$. The distance metric based on the energy associated with the flow through the tangent space is based on the Hilbert-Schmidt norm of an $n \times n$ matrix to be $\|A\|_{HS}^2 = tr AA^t$.

Definition 4 *The distance $\rho(\cdot, \cdot)$ between curves $x, x' \in \mathcal{M}(1)$ the space of smooth regular curves becomes*

$$\rho(x, x') \;=\; inf_{\phi \in \Phi : \phi \circ x = x'} \rho(x, x'; \phi) \tag{120}$$

$$\text{where } \rho(x, x'; \phi) \;\doteq\; \int_0^1 tr \underbrace{(F^x(s) - F^{x'}(s))(F^x(s) - F^{x'}(s))^t}_{\|F^x - F^{x'}\|_{HS}^2} ds \tag{121}$$

$$=\; b\int_0^1 |\alpha^x(s)\kappa^x(s) - \alpha^{x'}(s)\kappa^{x'}(s)|^2 ds + c\int_0^1 |\alpha^x(s)\tau^x(s) - \alpha^{x'}(s)\tau^{x'}(s)|^2 ds .$$

This involves generating all possible diffeomorphisms mapping one curve to another and choosing the minimizer of the above equation. This forms the basis for the curve matching. Define $\| \cdot \|_{\mathcal{V}}^2$ to be the norm-squared on velocity functions indexed over the curves.

Problem Statement 3 (Curve Matching) *Given are curves x, x' with skew-symmetric fields $F(\cdot), F'(\cdot)$, 3×3 skew-symmetric matrices encoding torsion-curvatures. The optimal match $\hat{h}$ of x to observation x' becomes*

$$\hat{h}(x, T) \;=\; \int_0^T -\nabla_x^t h(x, t)\hat{v}(x, t)dt + x , \tag{122}$$

$$\hat{v}(\cdot) \;\doteq\; \arg\min_{v \in \mathcal{V}} \|v\|_{\mathcal{V}}^2 + \int_0^1 \|F(h(s)) - F'(s)\|_{HS}^2 ds . \tag{123}$$

Bipartite graph matching. For matching curves, we have not yet implemented Eqn. 123, instead we follow the work of Bakircioglu [25] in which the cost function is associated with a bipartite graph which may be efficiently minimized over the space of discrete matches. For this, the discrete distance ρ^Δ associated with discrete curves on the triangulated surface is defined assuming the curves have piecewise constant curvature and torsion over segments:

$$\rho^\Delta(x, x'; \phi) \doteq b \sum_k |\alpha_k^x \kappa_k^x - \alpha_k^{x'} \kappa_k^{x'}|^2 \Delta_k + c \sum_k |\alpha_k^x \tau_k^x - \alpha_k^{x'} \tau_k^{x'}|^2 \Delta_k .$$

In [25] Bakircioglu et al. has used a smoothness term that depends only on the velocity giving a discrete matching cost according to

$$\hat{\phi} \doteq \arg\min_{\phi \in \Phi} a \sum_k |\alpha_k^x - \alpha_k^{x'}|^2 + b \sum_k |\alpha_k^x \kappa_k^x - \alpha_k^{x'} \kappa_k^{x'}|^2 \Delta_k + c \sum_k |\alpha_k^x \tau_k^x - \alpha_k^{x'} \tau_k^{x'}|^2 \Delta_k . \tag{124}$$

The discrete matching problem is solved by searching over all correspondences which minimize Eqn.124. Shown in Figure 21 are results of applying the bipartite matching to the Superior Temporal sulcus of two macaque whole brains measured in the laboratory of David Van Essen. Panel 1 shows the match based on penalizing the speed, $a = 1, b = 0, c = 0$; panel 2 shows the match based on curvature $a = 0, b = 1, c = 0$; panel 3 shows the match based on torsion, $a = 0, b = 0, c = 1$. The bottom row shows the same as above for the Arcuate sulcus of $87A$ matched to $90C$.

8.3 Probability Laws on Curves based on the Frenet Representation

Fundus curves in the macaque have been classified using standard nomenclature arising in [78, 135, 134]. Sulci exhibit pronounced variability in size and shape, motivating the development of representations of sulcus bed variability. There
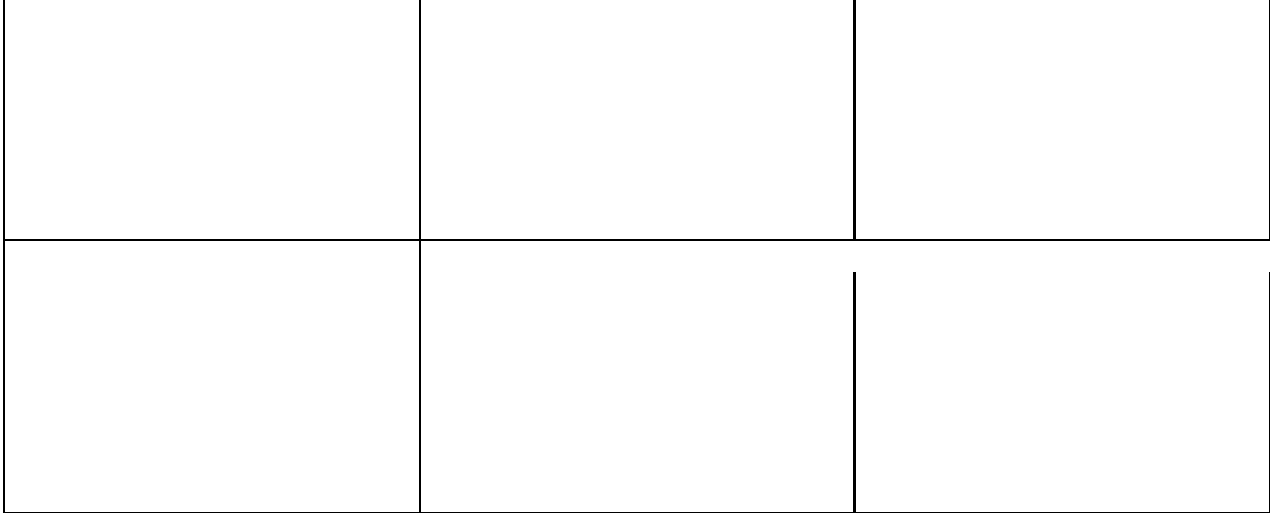
Figure 21: Top row: Figure shows the matching of the Superior Temporal sulcus of two macaque brains 87A and 93I measured in the laboratory of David Van Essen. Panel 1 shows speed matching for parameters speed $a = 1, b = 0, c = 0$, panel 2 shows curvature matching $a = 0, b = 1, c = 0$, panel c shows torsion matching $a = 0, b = 0, c = 1$. Bottom row: Shows the same as above for the Arcuate sulcus of $87A$ matched to $90C$.

are alternative representations available for describing fundus bed curves. Variability can be studied by inducing probabilities on the space of affine group transformations as in [70, 21]. Herein we induce probabilities using the intrinsic geometry of curves based on the Frenet characterization in torsion and curvature. To each curve $x(s)$ parametrized by arc length s associate the orthonormal frame field T, N, B. As per the Frenet representation, if $x(\cdot)$ is a unit speed curve with curvature and torsion fields κ, τ then $T'(s) = \kappa N(s)$, $N'(s) = -\kappa T(s) + \tau B(s)$, $B'(s) = -\tau N(s)$. For the prior distribution, associate the potential energy with the mean lengths, curvatures, and torsions $\bar{l}, \bar{\kappa}, \bar{\tau}$:

$$U(l, \kappa, \tau) = \eta \sum_k (l_k - \bar{l}_k)^2 + \beta \sum_k (\kappa_k - \bar{\kappa}_k)^2 + \gamma \sum_k (\tau_k - \bar{\tau}_k)^2 \ , \tag{125}$$

where η, β, γ are proportional to the total variance of the statistics around the mean. Shown in the left column of Figure 22 are two macaque brains with hand generated curves of the deep folds. From these the empirical estimators of the mean length, curvature and torsion functions of the fundus beds are computed for the $M = 3$ brains:

$$\hat{\bar{l}}_k = \frac{1}{M} \sum_{m=1}^{M} l_k^m, \quad \hat{\bar{\kappa}}_k = \frac{1}{M} \sum_{m=1}^{M} \kappa_k^m, \quad \hat{\bar{\tau}}_k = \frac{1}{M} \sum_{m=1}^{M} \tau_k^m \tag{126}$$

$$\hat{\eta}^{-1} = \frac{1}{MN} \sum_{m=1}^{M} \sum_k (l_k^m - \hat{\bar{l}}_k)^2, \quad \hat{\beta}^{-1} = \frac{1}{MN} \sum_{m=1}^{M} \sum_k (\kappa_k^m - \hat{\bar{\kappa}}_k)^2, \quad \hat{\gamma}^{-1} = \frac{1}{MN} \sum_{m=1}^{M} \sum_k (\tau_k^m - \hat{\bar{\tau}}_k)^2 \ .$$

To synthesize the fundus curves the random process $\{l_k, \kappa_k, \tau_k, k = 1, \cdots, N\}$ is generated to be Gaussian, with the discretized Frenet equations solved sequentially. Note the first two tangent vectors in the discrete representation are undetermined expressing the fact that the patterns are in the equivalence set modulo the special Euclidean group. Figure 22 shows results of such a probabilistic representation on fundi sampled to 50 points. The data for the fundus beds was obtained from three macaque brains labelled 87A, 90C and 93G. The statistics on torsion and curvature were computed. The middle column shows the mean fundus curves and varability around the mean curves. The top panel (middle column) shows the mean sulcus beds as displayed on the mean brain (see [21]). The bottom panel (middle column) shows sulcus beds synthesized from the Frenet equations driven by random curvatures and torsions sampled from the empirically estimated prior distribution.

Bayes estimation of the fundus curves. Incorporating the prior on the shape of fundi into the dynamic programming algorithm is straightforward. Define the curve $\alpha(s, t) \in P_{s,t}(S)$ to be a curve connecting points s, t on the triangulate graph passing through node points $x_k, k = 1, \ldots, N$ with associated lengths, curvatures and torsions
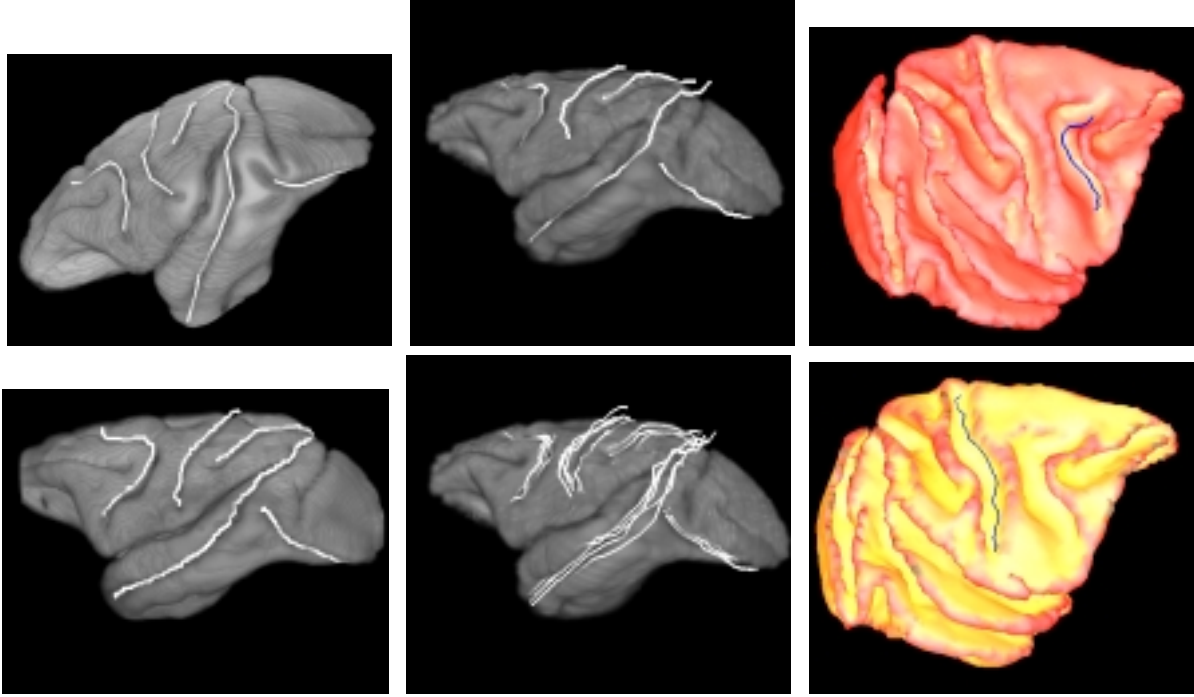
Figure 22: Left column shows macaque whole brains 90C and 93G with hand drawn fundus beds. The middle column shows the mean and variance results. Top row (middle column) shows the mean fundus beds computed from three macaque brains, 90C,93G, and 87A displayed on the mean brain; bottom row shows the fundus beds sampled from the prior distribution on curvature and torsion of the curve. Right column shows the Arcuate and Central fundus curves generated with the likelihood and the prior distribution. Data taken from the laboratory of David Van Essen of the Department of Anatomy and Neurobiology at Washington University.

$l_k, \kappa_k, \tau_k, k = 1, \ldots, N$. Define the posterior potential of a discrete curve $\alpha(s,t)$ connecting (s,t) as

$$
\begin{aligned}
H(\alpha) \quad = \quad & \eta \sum_k (l_k - \bar{l}_k)^2 + \beta \sum_k (\kappa_k - \bar{\kappa}_k)^2 + \gamma \sum_k (\tau_k - \bar{\tau}_k)^2 && (127) \\
+ \quad & \frac{1}{\sigma^2} \sum_k \left(\frac{(\kappa_{max}(x_k) - \mathcal{K})^2 + (\kappa_{max}(x_{k+1}) - \mathcal{K})^2}{2} + \frac{(\kappa_{max}(x_k) - \kappa_{max}(x_{k+1}))^2}{6} \right) \|x_{k+1} - x_k\| \ .
\end{aligned}
$$

Dynamic programming maximizes over the space of curves $\alpha(s,t) \in P_{s,t}(S)$.

Corollary 4 *Given s and t the start and end point of the curves $\alpha(s,t) \in P_{s,t}(S)$, then the algorithm in Theorem 8 with the costs for transition*

$$
\begin{aligned}
c^k(i,j) \quad = \quad & \infty, \ j \notin \mathcal{P}_i && (128) \\
c^k(i,j) \quad = \quad & \frac{1}{\sigma^2} \left(\frac{(\kappa_{max}(x_i) - \mathcal{K})^2 + (\kappa_{max}(x_j) - \mathcal{K})^2}{2} + \frac{(\kappa_{max}(x_i) - \kappa_{max}(x_j))^2}{6} \right) \|x_i - x_j\| && (129) \\
& + \eta(l_k - \bar{l}_k)^2 + \beta(\kappa_k - \bar{\kappa}_k)^2 + \gamma(\tau_k - \bar{\tau}_k)^2 \ , \ j \in \mathcal{P}_i \ , && (130)
\end{aligned}
$$

produces the maximum aposteriori estimator maximizing Eqn. 127, $\hat{\alpha}(s,t) \doteq \arg\min_{\alpha(s,t) \in P_{s,t}(S)} H(\alpha)$.

Shown in the right column, panels 3 and 6 of Figure 22 are results using of the prior distribution for generating the Arcuate and Central sulci, respectively, with $\hat{\bar{l}}, \hat{\bar{\kappa}}, \hat{\bar{\tau}}, \eta, \beta, \gamma$ estimated from the whole macaque brains 87A, 90C and 93G. Notice how smooth the curves are.

9 Symmetry in the Human Brain

The study of symmetry and lack of symmetry in the brain is important biologically. This can be made quantitatively precise. We have been studying the symmetry of the human brain via standard mathematical representations known as *symmetry groups* following the work of Lei Wang on the hippocampus. Most straightforwardly let us examine the *reflection group* **D** expressing axis flipping symmetry around the *sagittal plane*. This makes quantitative the notion of left and right sided brain symmetry, and variation away from left and right sided symmetry.

Begin by embedding the human brain in Ω the unit cube, with coordinates

$$x = (x_1, x_2, x_3); -1/2 \leq x_1 \leq 1/2; -1/2 \leq x_2 \leq 1/2; -1/2 \leq x_3 \leq 1/2 .$$

For symmetry around the *sagittal plane*, choose the convention that the sagittal plane P_{sag} through the center of the longitudinal cerebral fissure is the set of all points

$$P_{sag} = \{x = (x_1, x_2, x_3) \in \Omega \ : \ x_3 \ = 0\} .$$

The reflection subgroup $\mathbf{D} = \{I, R\}$ consists of the identity transformation I, the 3×3 identity matrix, and the reflection R around the sagittal plane becomes

$$\mathbf{D} = \{I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \ , \ R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \} . \tag{131}$$

The reflection generates the following global map around the sagittal plane:

$$R : (x_1, x_2, x_3) \quad \mapsto \quad (x_1, x_2, -x_3) .$$

Finding automatically the *sagital plane* around which the flip group is defined is being accomplished in our group as follows. We use the rigid motions to characterize lines and planes of symmetry. Define the rigid Euclidean transformation $(O, a) : x \mapsto y = Ox + a$, O an orthogonal matrix. Then the *plane of symmetry* is the set of points that are invariant to the above transformation, i.e. it is the plane defined by the points that satisfy the equation:

$$x = Ox + a .$$

Once the sagital plane is defined, then the question of right and left-sidedness, and its quantitative representation becomes straightforward. Examine the transformation from the left to right side of the brain, taking the left hand side (LHS) as the template, and the right hand side (RHS) as the target. The transformation is constrained to be of the form

$$h : x \mapsto h(x) = Rx - u(Rx) , \tag{132}$$

with R the reflection flip group. If the brain were perfectly symmetric, then $u(\cdot)$ is identically zero. The variation of $u(\cdot)$ away from zero is a measure of asymmetry. The average size and locality of variation of the $u(\cdot)$ fields from LHS to RHS then becomes a quantitative measure of asymmetry within the normal population and across the populations. Shown in Figure 23 are an example of the symmetries in the human brain between the left and right hippocampus. The left column shows a section through an MRI of the template hippocampus (top row) and a 3-D rendering of the template (bottom row). The middle column shows the LHS target that the template has been mapped into. The right column shows the result of flipping via the rigid axis-flipping group, and then mapping the LHS to the RHS.

10 Abnormal Anatomy

Now examine the generalization to abnormal anatomies. In formalizing pathologies we shall employ the dichotomy *automorphic pathologies* vs *heteromorphic pathologies* following the terminology in [2, 142]. An automorphic pathology modifies an anatomical configuration $I \in \mathcal{I}^\alpha$ into another $I' \in \mathcal{I}^\alpha$ via the diffeomorphism $h' \in \mathcal{H}$: $I' = h'I$, so that we do not leave the space of anatomies $\mathcal{I}^\alpha$. This is not true for the heteromorphic pathologies for which the new configuration is generated via a more drastic change of the regular topology, perhaps adding neoplasm or eliminating structures. This requires us to expand to multiple anatomies, $\mathcal{I} = \cup_\alpha \mathcal{I}^\alpha$, with the transformations acting $\mathcal{H} : \mathcal{I}^\alpha \to \mathcal{I}$.

Throughout we shall use the terms *normal/abnormal* in a way that differs from the customary one. Abnormal signifies something exceptional, an anomaly, but not necessarily indicating improper functioning. As well, throughout we shall concentrate solely on change in topological structure. There is of course a second type type of automorphic pathology, ones corresponding to changes in the textures of the normal textbook pictures $I \in \mathcal{I}$. We shall not explore this here. Throughout our discussions below we assume a two dimensional background space $\Omega \doteq [0, 1]^2$ in the small deformation setting.
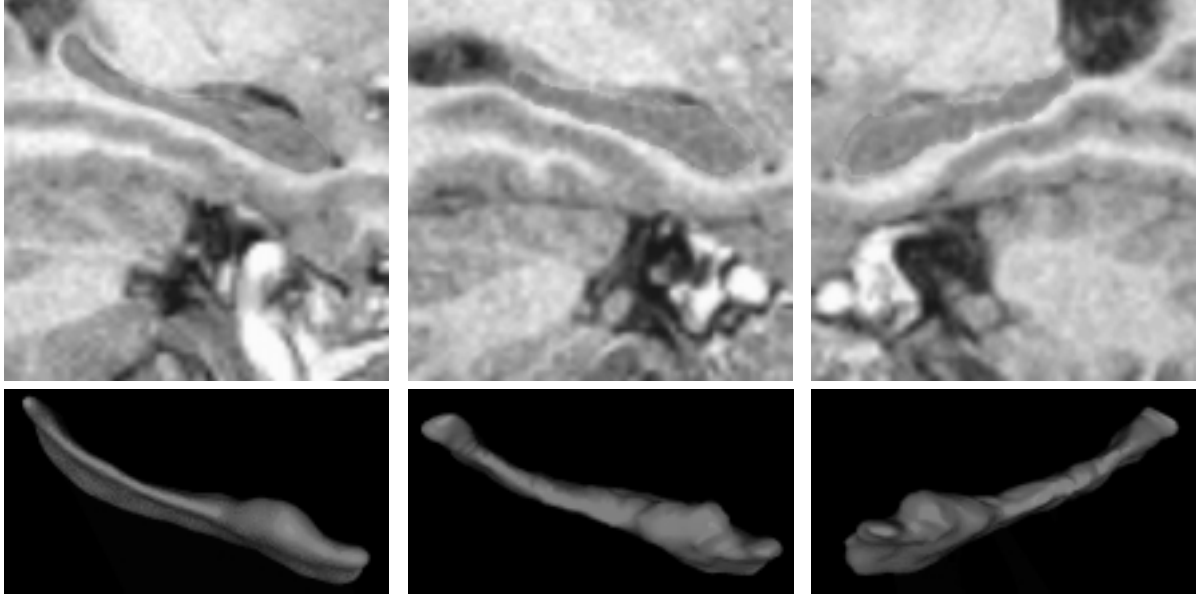
Figure 23: Left column shows the template section of MRI (top row) and 3-D surface rendering (bottom row). The middle and right panels show the template mapped to the RHS (middle column) and the RHS mapped to the LHS (right column).

10.1 Automorphic Pathologies

Automorphic pathologies, though abnormal, should *preserve the topological structures of a normal human body*. If the deformation operator $h \in \mathcal{H}$ is too far from the identity element in the similarity group so that it falls outside of a confidence set of the null hypothesis, H_0, we can identify this kind of abnormality by estimating the probability density $p(h)$ and applying the Neyman-Pearson lemma to arrive at an appropriate decision. Examining any standard atlas on disease, one is impressed with the vast number of mechanisms in the human body that can generate automorphic pathologies. Take arterial deformation where arteries have a decreased diameter. The decrease in diameter corresponds to h-values in a range where the density on deformations $p(h)$ is so small that it would occur only rarely for a normal patient. In the HANDS study [143] this idea of testing h-values far from the identity element in the group was used for detecting abnormalties such as a malformed thumb.

Examine automorphic shape changes. Say that a planar anatomy is shrinking or expanding. Then the group of transformation $\mathcal{H}$ take the form $\mathcal{H} : x = (x_1, x_2) \rightarrow h(x) = x - (u_1(x), u_2(x))$. Introduce the dilatation vector $d = (d_1, d_2) \in I\!\!R^2$. Abnormal anatomy H_1 is generated via the solution of the random equation $LU = W$, with (L^{H_0}, W^{H_0}) associated with normal, and (L^{H_1}, W^{H_1}) associated with disease.

The energy functions on the $N \times N$ lattice in the Gaussian prior distribution could take the form

$$E(u) = \sum_{x_1=1}^{N} \sum_{x_2=1}^{N} Q(x_1, x_2)$$

$$\text{where} \quad H_0 \quad : \quad Q(x_1, x_2) = (\Delta u_1)^2 + (\Delta u_2)^2; (x_1, x_2) \tag{133}$$

$$H_1 \quad : \quad Q(x_1, x_2) = (\Delta u_1 - d_1)^2 + (\Delta u_2 - d_2)^2; (x_1, x_2) \ , \tag{134}$$

and with Δ standing for the discrete Laplacian operator

$$\Delta u_1(x_1, x_2) \quad = \quad u_1(x_1+1, x_2) - 2u_1(x_1, x_2) + u_1(x_1-1, x_2) + u_1(x_1, x_2+1) - 2u_1(x_1, x_2) + u_1(x_1, x_2-1) \tag{135}$$

$$\Delta u_2(x_1, x_2) \quad = \quad u_2(x_1+1, x_2) - 2u_2(x_1, x_2) + u_2(x_1-1, x_2) + u_2(x_1, x_2+1) - 2u_2(x_1, x_2) + u_2(x_1, x_2-1) \tag{136}$$

Assume cyclic or fixed boundary conditions. To synthesize the prior measure under P_{H_0}, P_{H_1} we solve the stochastic partial difference equations

$$H_0 \quad : \quad (\Delta u_j)(x_1, x_2) = w_j(x_1, x_2) \ , \ j = 1, 2 \ , \tag{137}$$

$$H_1 \quad : \quad (\Delta u_j)(x_1, x_2) = w_j(x_1, x_2) + d_1 \ , \ j = 1, 2 \ , \tag{138}$$

where $w(\cdot, \cdot)$ is a noise realization. A slight modification of this would be to allow the vectors d_1, d_2 to be space dependent, for example pointing outwards or inwards from a center of the abnormality.

In the left panel of Figure 24 the deformations are shown on an MRI section from a human brain for various choices of the dilatation vector. The top left panel a) shows the original image with the area of dilatation depicted; the top right panel b) shows a contracting field $d1 < 0$ and $d2 < 0$; the bottom left panel c) shows an expanding field $d1 > 0$ and $d2 > 0$; the bottom right panel d) shows a shearing field $d1 < 0$ and $d2 > 0$. The right column shows similar results for the ventricles.
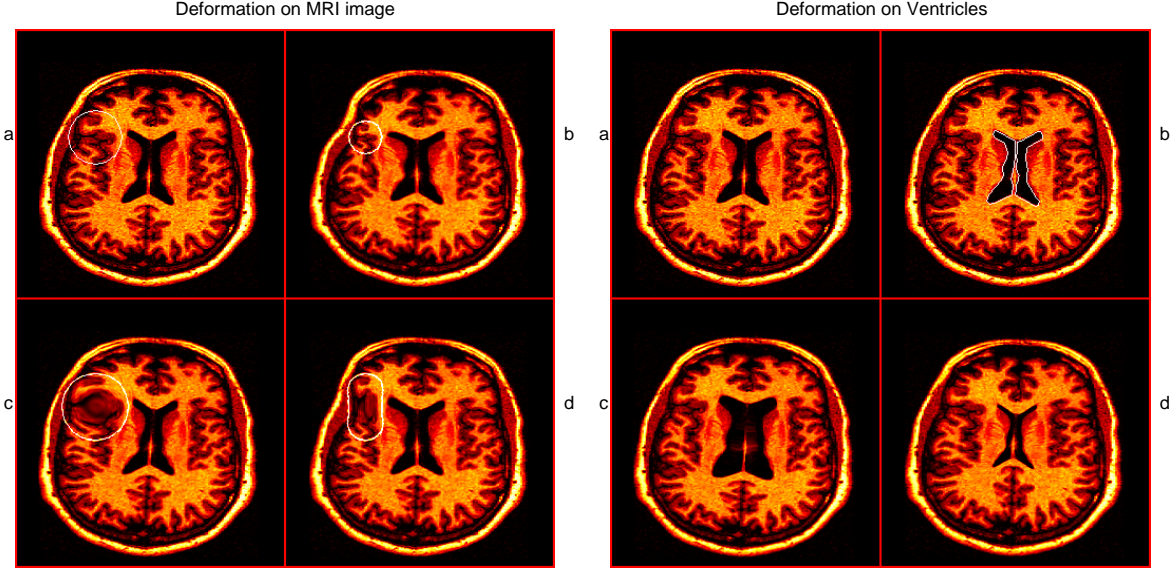

Figure 24: Figures showing homomorphic deformations. Left figure: Top left panel a) shows the original image with the area of dilatation depicted; the top right panel b) shows a contracting field $d1 < 0$ and $d2 < 0$; the bottom left panel c) shows an expanding field $d1 > 0$ and $d2 > 0$; the bottom right panel d) shows a shearing field $d1 < 0$ and $d2 > 0$. Right figure: Shows analogous deformations to the ventricles.

Testing for automorphic abnormality. Under the null (normal) and disease hypotheses let the random deformation have densities $H_0 : p_0(h)$, $H_1 : p_1(h)$. Assume the image I^D are generated of a particular patient's anatomy, the goal being to diagnose disease. The likelihood function of the measured imaged data I^D expresses the physical chararcteristics of the sensor(s) used and therefore does not depend upon the choice of prior, thus $L(I^D|h)$ is the same for both hypotheses H_0, H_1. This implies that the density $f_0(\cdot)$ of the observed deformed image I^D under the hypotheses H_0, H_1 becomes

$$H_0 \ : \ f_0(I^D) \ = \ \int_{\mathcal{H}} p_0(h) L(I^D, h) dh \ ; \tag{139}$$

$$H_1 \ : \ f_1(I^D) \ = \ \int_{\mathcal{H}} p_1(h) L(I^D, h) dh \tag{140}$$

The question of how to calculate such integrals has been addressed in [2]. The most powerful test comes out directly from the Neyman-Pearson lemma in terms of the critical region W as

$$W = \{I^D : \frac{f_1(I^D)}{f_0(I^D)} \geq constant\} \ .$$

If the sensor is very accurate, which should be interpreted as meaning that the posterior densities $f_0 \propto p_0(\cdot) L(I^D, \cdot)$, $f_1 \propto p_0(\cdot) L(I^D, \cdot)$ are very peaked, then the group element $h \in \mathcal{H}$ is effectively observed with high accuracy, so that the test can be organized in terms of the group elements directly. This is probably true for high quality MRI's, and certainly true for CRYOSECTION and post-mortem imagery. Then the most powerful test effectively becomes

$$W = \{h : \frac{p_1(h)}{p_0(h)} \geq constant\} \ .$$

10.2   Heteromorphic pathologies

Dealing with heteromorphic pathologies is a more difficult challenge. Clearly new masses due to various tumors introduce *deformation which carry us outside* the space $\mathcal{I}^\alpha$, requiring the introduction of a larger sample space $\mathcal{I} = \cup_\alpha \mathcal{I}^\alpha$, with $\mathcal{H} : \mathcal{I}^\alpha \to \mathcal{I}$. To represent this type of heteromorphic pathology, where *new mass pushes out the normal tissue*, we shall use a deformable template such as in [7] for the mass and then condition the h-field accordingly.

Introduce a template configuration of the domain $\Omega^{path} \subset \Omega$ representing the mass and denote its central point by x_{center}. On this template let the random transformations act. To fix ideas let the translation groups act on the background spaces $\Omega_{path} \subset \Omega$, and use a stochastic difference equation $LU = E$ to describe the normal displacements in $\Omega \setminus \Omega^{path}$. The normal tissue around x_{center} is displaced by the new mass.

Denote the boundary of the pathology Ω^{path} as $\partial \Omega^{path}$, and solve the same stochastic difference equation as before, but in the domain outside of the new mass

$$(LU)(x) = W(x), x \in \Omega \setminus \Omega^{path} \ ,$$

with the boundary conditions $U(x) = x_{center} - x, x \in \partial \Omega^{path}$. In other words, the normal material is pushed out from x_{center} and replaced by the tissue in Ω^{path}.

The left panel of Figure 25 shows an expansion or "push out"; the right panel shows a "push in". Another pathology
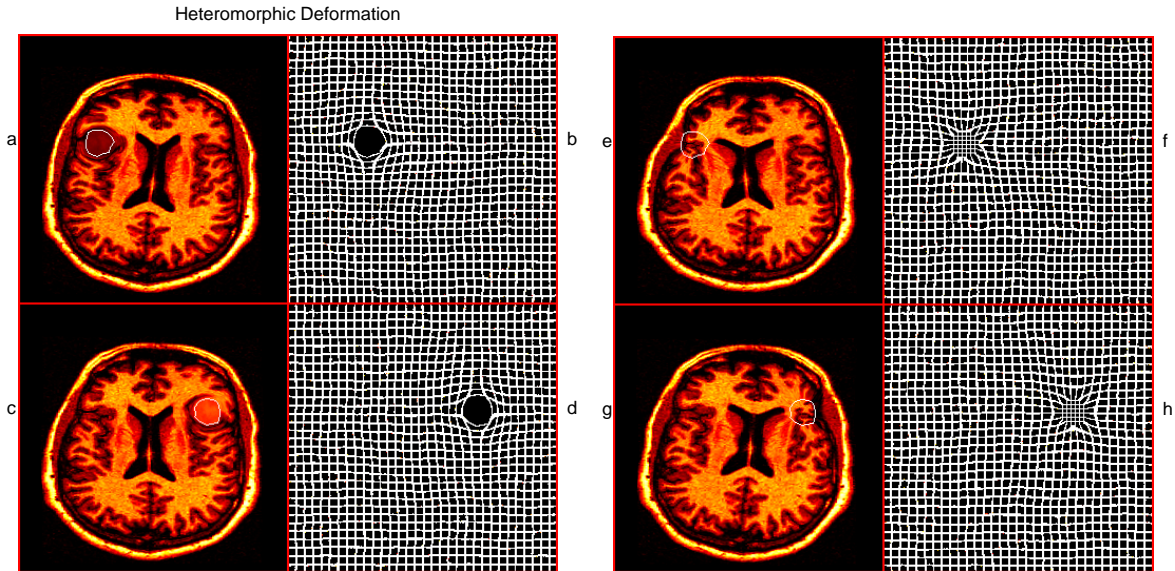

Figure 25: Figure showing heteromorphic deformations. The left panel shows an expansion or "push out"; the right panel shows a "push in".

that may be of interest is when the change in tissue is not a "push out" but "cover" , in the sense that the new mass takes the place of some of the normal ones, invades it without displacing it. This is straightforwardly represented by solving the stochastic difference equation over the whole space Ω, and then placing a deformed template at some x_{center} covering the rest in that part of Ω.

To represent heteromorphic shape abnormalities assume that the anatomy is enclosed in a bounded rectangular background space $X \subset I\!R^3$ and use the operator $L = \Delta + c, c < 0$, with Dirichlet boundary conditions on $\partial \Omega$. Making c negative ensures non-singularity and existence of a solution to the problem below. For simplicity we assume the three components of the deformation field $d(x); x \in \Omega$ to be independent; for the coupled case we should replace L by the Navier operator used above.

Let $a \in \Omega$ be a center for an expanding abnormality that expands into a region $A \subset \Omega$. For example, A could be a ball with the center at a but more general shapes will be allowed. The *normal* remainder will be denoted $N = \Omega \setminus A$. With the usual representation for the observed image $I^D(x) \doteq I_{temp}[x - u(x)]$ so that we are led to the following stochastic field equation for the pathology induced displacement

$$(Lu)(x) \ = \ (\Delta u)(x) \ + \ cu(x) \ = \ w(x) \ , \ x \in N \ ,$$

with the boundary condition

$$u(x) \ = \ u_A(x) \ = \ a \ - \ x; x \in \partial A$$

$$u(x) \ = \ 0; x \in \partial\Omega$$

and $w(\cdot)$ is a noise field. Note that $u(\cdot)$ is not and should not be defined in the abnormal region A.

Let $G = G(x, y); x, y \in N$ be the Green's function for the Laplacian Δ in N with Dirichlet conditions. In N introduce a function $v(x)$ taking the values $u_A(x)$ on ∂A and zero on $\partial\Omega$. For the new field $g = u - v$ we have

$$Lg \ = \ Lu - Lv \ = \ w \ - \ Lv$$

and g satisfies Dirichlet boundary conditions. Ansatz:

$$g(x) \ = \ \int_N G(x, y) f(y) dy$$

with some function f to be determined. Then

$$(Lg)(x) \ = \ \Delta \int_N G(x, y) f(y) dy \ + \ c \int_N G(x, y) f(y) dy \ = f(x) \ + \ \int_A G(x, y) f(y) dy$$

But $Lg = Lu - Lv = w - Lv$ so that we get the integral equation

$$f(x) \ + \ \int_N G(x, y) f(y) dy \ = \ g(x) \ - (Lv)(x)$$

and we should solve the Fredholm equation with the Fredholm alternative holding.

Remark 10 *For a more complete simulation we must define the texture inside A. Because of the heteromorphic property this texture is not inherited from the template but must be defined separately. Methods such as that used in [7] are then appropriate.*

Remark 11 Random abnormality region. *In the above we assumed A to be given in advance; an alternative would be to let it grow randomly, by introducing a force field $F(x); x \in \Omega$ taking 3-vectors as values and such that is directed outward from the abnormality center a and solve*

$$(Lu)(x) \ = \ w(x) + F(x); x \in \Omega \ .$$

11   Conclusions

Within this paper we have developed mathematical representations of the variability of human anatomy focussing on neuroanatomical variation of neocortical volumes, and 1,2-dimensional sulcal and surface submanifolds. The generation of probabilistic measures of anatomical variation on coordinatized submanifolds is formulated as an empirical procedure in which populations of brains are mapped to common coordinate systems. Coordinate systems are constructed which are closest to the population of anatomies in a minimum distance sense, the idealized coordinate systems become the empirically estimated templates.

High dimensional Gaussian probability measures on vector field deformations of the template coordinates and its submanifolds are defined. Covariance estimation is posed as generalized spectrum estimation indexed over the submanifolds, with parameters of the spectral representations estimated directly from the empirically measured maps onto the anatomical populations. Mathematical methods for estimating brain volumes as well as manifold variations associated with closed smooth surface representations associated with hippocampal subvolumes, and cortical sulci have been described.

Several methods are described showing the use of these mathematical representations of variability for characterizing *disease* described as classical Bayesian hypothesis testing against base measures. *Disease* or *abnormality* refers to either (1) local and regional alterations in the size and shape of brain structures relative to the base normal measure, i.e. a *large deviation*, or (2) the abnormal presence or absence of substructures corresponding to a fundamental change in topology of the underlying graph structure of the brain.

The mathematical methods applied in this paper may be unfamiliar to many researchs in computational anatomy. We believe, however, that they the methodology presented herein will form the basis for studying complex anatomical structures.

12   Acknowledgements

References

[1] U. Grenander. A unified approach to pattern analysis. *Adv. Computing.*, 10:175–216, 1970.

[2] U. Grenander. *General Pattern Theory.* Oxford Univ. Press, 1994.

[3] W. M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry.* Academic Press, 1986.

[4] M.I. Miller, G.E. Christensen, Y. Amit, and U. Grenander. Mathematical textbook of deformable neuroanatomies. *Proceedings of the National Academy of Science*, 90(24), December 1993.

[5] G. E. Christensen, R. D. Rabbitt, and M.I. Miller. A deformable neuroanatomy textbook based on viscous fluid mechanics. In Jerry Prince and Thordur Runolfsson, editors, *Proceedings of the Twenty-Seventh Annual Conference on Information Sciences and Systems*, pages 211–216, Baltimore, Maryland, March 24-26 1993. Department of Electrical Engineering, The Johns Hopkins University.

[6] G. E. Christensen, R. D. Rabbitt, and M.I. Miller. 3D brain mapping using a deformable neuroanatomy. *Physics in Medicine and Biology*, 39:609–618, 1994.

[7] U. Grenander and M. I. Miller. Representations of knowledge in complex systems. *Journal of the Royal Statistical Society B*, 56(3):549–603, 1994.

[8] M. W. Vannier, M. I. Miller, and U. Grenander. Modeling and data structure for registration to a brain atlas of multimodality images. In edited by R. W. Thatcher, M. Hallett, T. Zeffiro, E. R. John, and M. Huerta, editors, *Functional Neuroimaging- Technical Foundations*, pages 217–221. Academic Press, 1994.

[9] Sarang C. Joshi, Michael I. Miller, Gary E. Christensen, Ayananshu Banerjee, Thomas A. Coogan, and Ulf Grenander. Hierarchical brain mapping via a generalized dirichlet solution for mapping brain manifolds. In *Proc. of the SPIE's 1995 International Symposium on Optical Science, Engineering, and Instrumentation*, volume 2573, pages 278–289, August, 1995.

[10] Sarang C. Joshi, Jing Wang, Michael I. Miller, David Van Essen, and Ulf Grenander. On the differential geometry of the cortical surface. In *Proceedings of SPIE's 1995 Geometric Methods in Applied Imaging*, San Diego, CA., 9-14 July, 1995.

[11] G. E. Christensen, R. D. Rabbitt, M.I. Miller, S.C. Joshi, U. Grenander, and T. A. Coogan. *Topological Properties of Smooth Anatomic Maps.* Kluwer Academic Publishers, 1995.

[12] G.E. Christensen, M.I. Miller, M. W. Vannier, and U. Grenander. Individualizing neuroanatomical atlases using a massively parallel computer. *Computer*, January 1996.

[13] G. E. Christensen, R. D. Rabbitt, and M. I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1435–1447, October 1996.

[14] John W. Haller, Gary E. Christensen, Sarang Joshi, , John W. Newcomer, Michael I. Miller, John C. Csernansky, and Michael W. Vannier. Hippocampal mr imaging morphometry by means of general pattern matching. *Radiology*, 199(3):787–791, June 1996.

[15] John W. Haller, Ayananshu Banerjee, Gary E. Christensen, Sarang Joshi, Michael I. Miller, Michael W. Vannier, and John C. Csernansky. Three-dimensional hippocampal volumetry by high dimensional transformation of a neuroanatomical atlas. *Radiology*, 202(2):504–510, February 1997.

[16] S. Joshi. *Large Deformation Diffeomorphisms and Gaussian Random Fields for Statistical Characterization of Brain SubManifolds.* PhD Thesis, Department of Electrical Engineering, Sever Institute of Technology, Washington University, St. Louis,Mo., August 1997.

[17] L. Matejic. *Group Cascades for Representing Biological Variability.* Ph.D. Dissertation and Brown University, Providence, RI, 1997.

[18] P. Dupuis, U. Grenander, and M.I. Miller. Variational problems on flows of diffeomorphisms for image matching. Report of the lcds, 1997.

[19] G. E. Christensen, S. C. Joshi, and M. I. Miller. Volumetric transformation of brain anatomy. *IEEE Transactions on Medical Imaging*, 16(6), December 1997.

[20] S. Joshi, M.I. Miller, and U. Grenander. On the geometry and shape of brain sub-manifolds. *International Journal of Pattern Recognition and Artificial Intelligence*, 11(8), 1997.

[21] M. Miller, Ayananshu Banerjee, Gary Christensen, Sarang Joshi, Navin Khaneja, U. Grenander, and Larissa Matejic. Statistical methods in computational anatomy. *Statistical Methods in Medical Research*, (6):267–299, 1997.

[22] P. Dupuis, U. Grenander, and M.I. Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*, 1998.

[23] D. C. Van Essen, H. Drury, S. Joshi, and M. I. Miller. Functional and structural mapping of human cerebral cortex: Solutions are in the surfaces. *Proceedings of the National Academy of Science*, 95:788–795, February 1998.

[24] M.I. Miller N. Khaneja, U. Grenander. Dynamic programming generation of geodesics and sulci on brain surfaces. *Pattern Analysis and Machine Intelligence*.

[25] Muge Bakircioglu, Ulf Grenander, Navin Khaneja, and Michael Miller. Curve matching on brain surfaces using frenet distances. *To appear in Human Brain Mapping*, 1998.

[26] C. M. Pechura and J. B. Martin. Mapping the brain and its functions: Integrating enabling technologies in neuroscience research. Report of the institute of medicine, 1991.

[27] Greitz, Bohm, Holte, and Eriksson. A computerized brain atlas: Construction, anatomical content, and some applications. *Journal of Computer Assisted Tomography*, 15(1):26–38, January/February 1991.

[28] R. Bajcsy, R. Lieberson, and M. Reivich. A Computerized System for the Elastic Matching of Deformed RAdiographic Images to Idealized Atlas Images. *Journal of Computer Assisted Tomography*, 7(4):618–625, 1983.

[29] R. Dann, J. Hoford, S. Kovacic, M. Reivich, and R. Bajcsy. Evaluation of Elastic Matching Systems for Anatomic (CT, MR) and Functional (PET) Cerebral Images. *Journal of Computer Assisted Tomography*, 13(4):603–611, July/August 1989.

[30] K.H. Hohne, M. Bomans, M. Reimer, R. Schubert, U. Tiede, and W. Lierse. A volume-based anatomical atlas. *IEEE Comput. Graphics Appl.*, pages 72–78, 1992.

[31] J.C. Mazziotta, Arthur W. Toga, Alan Evans, Peter Fox, and Jack Lancaster. Probabilistic atlas of the human brain: Theory and rationale for its development. *Neuroimage*, 2:89–101, 1995.

[32] M.E. Shenton, R. Kikinis, F.A. Jolesz, S.D. Pollak, M. LeMay, C.G. Wible, H. Hokama, J. Martin, D. Metcalf, M. Coleman, and R. W. McCarley. Abnormalities of the left temporal lobe and thought disorder in schizophrenia. *N. Engl. J. Med.*, 327:604–612, 1992.

[33] W. Jansen, J.P. Baak, A.W. Smeulder, and A.M. van Ginneken. A computer based handbook and atlas of pathology. *Pathol. Res. Pract.*, 185(5):652–656, 1989.

[34] C. Bohm, T. Greitz, B. Berggren, and L. Ollson. Adjustable computerized stereotaxic brain atlas for transmission and emission tomography. *AJNR*, 4:731–733, 1988.

[35] *Extraction of morphometric information from dual echo magnetic resonance brain images*, volume 1360, 1991.

[36] A.C. Evans, C. Beil, S. Marret, C.J. Thompson, and A. Hakim. Anatomical-functional correlation using an adjustable mri-based region of interest atlas with positron emission tomography. *Journal of Cerebral Blood Flow and Metabolism*, 8:513–530, 1988.

[37] A.C. Evans, W. Dai, L. Collins, P. Neelin, and S. Marret. Warping of a computerized 3-d atlas to match brain image volumes for quantitative neuroanatomical and functional analysis. *Image Processing*, 1445:236–246, 1991.

[38] D. C. Van Essen and H. Drury. Structural and functional analyses of human cerebral cortex using a surface-based atlas. *J. Neuroscience*, in press.

[39] F.L. Bookstein. *The Measurement of Biological Shape and Shape Change*, volume 24. Springer-Verlag: Lecture Notes in Biomathematics, New York, 1978.

[40] F.L. Bookstein and W.D.K. Green. Edge information at landmarks in medical images. In Richard A. Robb, editor, *Visualization in Biomedical Computing 1992*, pages 242–258. SPIE 1808, 1992.

[41] F.L. Bookstein. Biometrics, biomathematics and the morphometric synthesis. *Bulletin of Mathematical Biology*, 58(2):313–365, 1996.

[42] A.W. Toga, P.K. Banerjee, and B. A. Payne. Brain warping and averaging. *J. Cereb. Blood Flow Metab.*, 11:S560, 1991.

[43] W.D.K. Green and F.L. Bookstein. Edge information at landmarks in medical images. In Richard A. Robb, editor, *Visualization in Biomedical Computing*, pages 242–258. SPIE, 1996.

[44] D. Terzopoulos. *Multiresolution computation of visible-surface representations*. Ph.D Thesis, Massachusetts Institute of Technology, Boston, MA., 1984.

[45] M. Kass, A. Witkin, and D. Terzopolous. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.

[46] D. Terzopoulos and K. Waters. Physically-based facial modelling, analysis, and animation. *The Journal of Visualization and Computer Animation*, 1:73–80, 1990.

[47] I. Carlbom, D. Terzopoulos, and K. Harris. Computer-assisted registration, segmentation, and 3d reconstruction from images of neuronal tissue sections. *IEEE Transactions on Medical Imaging*, 13(2):351–362, june 1994.

[48] D. Terzopoulos and D. Metaxas. Dynamic 3d models with local and global deformations: Deformable superquadrics. *IEEE Trans. Patt. Anal. Mach. Intell.*, 13:703–714, 1991.

[49] A. Pentland and S. Sclaroff. Closed-form solutions for physically based shape modeling and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):715–729, July 1991.

[50] L.H. Staib and J.S. Duncan. Boundary finding with parametrically deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-14(11):1061–1075, November 1992.

[51] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Active shape models—their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.

[52] M.I. Miller, S. Joshi, D. R. Maffitt, J. G. McNally, and U. Grenander. Mitochondria, membranes and amoebae: 1,2 and 3 dimensional shape models. In Kanti Mardia, editor, *Statistics and Imaging*, volume II. Carfax Publishing C., Abingdon, Oxfordshire - England, 1994.

[53] K.J. Friston, C.D. Frith, P.F. Liddle, and R.S.J. Frackowiak. Plastic transformation of pet images. *Journal of Computer Assisted Tomography*, 15:634–639, 1991.

[54] R.D. Rabbitt, J.A. Weiss, G.E. Christensen, and M.I. Miller. Mapping of hyperelastic deformable templates using the finite element method. *Presented at the International Symposium on Optical Science, Engineering and Instrumentation*, July 1995.

[55] John W. Haller, Michael I. Miller Gary E. Christensen, Sarang Joshi, Mokhtar Gado, John Csernansky, and Michael W. Vannier. A comparison of auatomated and manual segmentation of hippocampus mr images. In *Proceedings of SPIE's Medical Imaging 1995*, volume 2434, pages 206–215, San Diego, California, 27February - 2 March,1995.

[56] John W. Haller, Gary E. Christensen, Sarang Joshi, Michael I. Miller, and Michael W. Vannier. Digital atlas-based segmentation of the hippocampus. In C.C. Jaffe H.U. Lemke, K. Inamura and M.W. Vannier, editors, *Computer Assisted Radiology:Proceedings of the International Symposium on Computer and Communications Systems for Image Guided Diagnosis and Therapy*, volume 2434, pages 152–157, CAR'95 Berlin, June 21-24, 1995.

[57] D.L. Collins, Peter Neelin, Terrence M. Peters, and A.C. Evans. Automatic 3d intersubject registration of mr volumetric data in standardized talairach space. *Journal of Computer Assisted Tomography*, pages 192–205, March/April 1994.

[58] K.J. Friston, J. Ashburner, C.D. Frith, J.-B. Poline, J. D. Heather, Liddle, and R.S.J. Frackowiak. Spatial registration and normalization of images. *Human Brain Mapping*, 2:165–189, 1995.

[59] D.L. Collins, C.J. Holmes, Terrence M. Peters, and A.C. Evans. Automatic 3-d model-based neuroanatomical segmentation. *Human Brain Mapping*, 3:190–208, 1995.

[60] C. Davatzikos, Marc Vaillant, Susan m. Resnick, Jerry L. Prince, Stanley Letovsky, and R. Nick Bryan. A computerized approach for morphological analysis of the corpus callosum. *Journal of Computer Assisted Tomography*, 20(1):88–97, 1996.

[61] C. Davatzikos. Spatial normalization of 3-d brain images using deformable models. *Journal of Computer Assisted Tomography*, 20(4):656,665, July/August 1996.

[62] C. Davatzikos. Spatial transformation and registration of brain images using elastically deformable models. *Comp. Vision and Image Understanding*, 66(2):207–222, May 1997.

[63] A. Dale and M. Sereno. Improved localization of cortical activity by combining eeg and meg with mri cortical surfaec reconstruction: A linear approach. *J. Cognit. Neuroscience*, 5:2:162–176, 1993.

[64] Rademacher J, Caviness Jr VS, Steinmetz H, and Galaburda AM. Topographical variation of the human primary cortices:implications for neuroimaging,brain mapping and neurobiology. *Cerebral Cortex*, 2:313–329, 1993.

[65] H. A. Drury, D. C. Van Essen, C. H. Anderson, C. H . Lee, T. A. Coogan, and J. W. Lewis. Computerized mappings of the cerebral cortex. a multiresolu tion flattening method and a surface-based coordinate system. *J. Cogn. Neurosci.*, 8:1–28, 1996.

[66] D.C. Van Essen. Pulling strings to build a better brain: A tension-based theory of morphogenesis and compact wiring in the central nervous sytem. *Nature*, 385:313–318, 1997.

[67] J.P. Thirion and A. Gourdon. 3d marching line algorithm and its applications to crest lines extraction. *Research Report No. 1672, INRIA, Epidaure*, 1992.

[68] J.P. Thirion and A. Gourdon. The 3d marching lines algorithm. *CVGIP: Graphical Models and Image Processing*, pages 503–509, 1996.

[69] J.T. Kent, K.V. Mardia, and J.M. West. Ridge curves and shape analysis. In *Monograph, Department of Statistics, University of Leeds.*, Leeds LS2 9JT, UK, May, 1996.

[70] N. Khaneja. *Statistics and Geometry of Cortical Features*. M.S. Thesis, Department of Electrical Engineering, Sever Institute of Technology, Washington University, St. Louis, MO, December 1996.

[71] L. Younes. Computable elastic distances between shapes. *To appear in SIAM J. Applied Math.*, 1998.

[72] M. Piccioni, S. Scarlatti, and A. Trouvé. A variational problem arising from speech recognition. *SIAM Jour. of Applied Maths*, 1998.

[73] A. Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *Intern. Jour. of Computer Vision.*

[74] A. Trouvé. Action de groupe de dimension infinie et reconnaissance de formes. C. R. Acad. Sci. Paris, Série I, ((321)):1031–1034, 1995.

[75] A. Trouvé. An infinite dimensional group approach for physics based models in patterns recognition.

[76] A. Trouvé. *habilitation à diriger les recherches*. University Orsay, *January 1996*.

[77] *S.K. Kyriacou, C. Davatzikos, S.J. Zinreich, and R.N. Bryan. Modeling brain pathology and tissue deformation using a finite element based nonlinear elastic models. IEEE Trans. on Med. Imag., 1997. submitted.*

[78] *D.J. Felleman and D.C. Van Essen. Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex, 1(1):1–47, 1991.*

[79] *L.D. Griffin. The Intrinsic Geometry of the Cerebral Cortex. Journal of Theoretical Biology, 166:261–273, 1994.*

[80] *L. A. Shepp and Y. Vardi. Maximum-likelihood reconstruction for emission tomography. IEEE Trans. on Medical Imaging, MI-1:113–121, 1982.*

[81] *L.A. Shepp, Y. Vardi, J. B. Ra, S.K. Hilal, and Z.H. Cho. Maximum-likelihood with real data. IEEE Transaction on Nuclear Science, NS-31:910–913, 1984.*

[82] *E. Veklerov and J. Llacer. Stopping rule for the mle algorithm based on statistical hypothesis testing. IEEE Trans. on Medical Imaging, MI-6, No. 4:313–319, 1987.*

[83] J. Llacer and E. Veklerov. *Feasible images and practical stopping rules for iterative algorithms in emission tomography.* IEEE Trans. on Medical Imaging, *MI-8:186–193, 1989.*

[84] M. I. Miller, D. L. Snyder, and T. R. Miller. *Maximum likelihood reconstruction for single photon emission computed tomography.* IEEE Trans. on Nuclear Science, *NS-32,No.1:769–778, 1985.*

[85] C.E. Floyd, R.J. Jaszczak, and R.E. Coleman. *Inverse monte-carlo: a unified reconstruction algorithm for SPECT.* IEEE Transactions on Nuclear Science, *NS-32:779–785, 1985.*

[86] G. T. Herman and D. Odhner. *Performance evaluation of an iterative image reconstruction algorithm for positron emission tomography.* IEEE Transactions on Medical Imaging, *MI-10:336–346, September 1991.*

[87] Jeffrey A. Fessler. *Hidden-data spaces for maximum-likelihood PET reconstruction.* In Conference Record of the 1992 IEEE Nuclear Science Symposium and Medical Imaging Conference, *pages 898–900, Orlando, Florida, October 25-31, 1992. IEEE Service Center.*

[88] J.-S. Liow and S. C. Strother. *Noise and signal decoupling in maximum likelihood reconstructions and metz filters for pet images.* Conference Record of the 1992 IEEE Nuclear Science Symposium and Medical Imaging Conference, *2:901–903, 1992.*

[89] D. G. Politte and D. L. Snyder. *Corrections for accidental coincidences and attenuation in maximum-likelihood image reconstruction for positron-emission tomography.* IEEE Transactions on Medical Imaging, *10(1):82–89, March 1991.*

[90] U. Grenander. Abstract Inference. *John Wiley and Sons, New York, 1981.*

[91] D.L. Snyder and M.I. Miller. *The use of sieves to stabilize images produced with the EM algorithms for emission tomography.* IEEE Trans. on Nuclear Science, *NS-32:3864–3872, 1985.*

[92] S. Geman and D.E. McClure. *Bayesian image analysis: An application to single photon emission tomography.* Proceedings of the American Statistical Association, *pages 12–18, 1985.*

[93] D.L. Snyder, M.I. Miller, Jr. L.J. Thomas, and D.G. Politte. *Noise and edge artifacts in maximum-likelihood reconstruction for emission tomography.* IEEE Transactions on Medical Imaging, *MI-6, No.3:228–237, 1987.*

[94] M.I. Miller and B. Roysam. *Bayesian image reconstruction for emission tomography: Implementation of the EM algorithm and Good's roughness prior on massively parallel processors.* Proc. of the Natl. Acad. of Sci., *88:3223–3227, April 1991.*

[95] E.S. Chornoboy, C.J. Chen, M.I. Miller, T.R. Miller, and D.L. Snyder. *An evaluation of maximum likelihood reconstruction for SPECT.* IEEE Trans. on Medical Imaging, *9(1):99–110, 1990.*

[96] T. Hebert and R. Leahy. *A generalized EM algorithm for 3-d Bayesian reconstruction from poisson data using Gibbs priors.* IEEE Trans. on Medical Imaging, *MI-8, No. 2:194–202, June 1989.*

[97] P. J. Green. *On use of the EM algorithm for penalized likelihood estimation.* J. Royal Statistical Society, B, 52, *No. 3:443–452, 1990.*

[98] K. Lange. *Convergence of EM image reconstruction algorithms with Gibbs smoothing.* IEEE Trans. on Medical Imaging, *MI-9, No. 4:439–446, December 1990.*

[99] A.W. McCarthy and M.I. Miller. *Maximum likelihood SPECT in clinical computation times using mesh-connected parallel computers.* IEEE Transactions on Medical Imaging, *10, No.3:426–436, September 1991.*

[100] G.T. Herman, A.R. De Pierro, and N. Gai. *On methods for maximum a posteriori image reconstruction with a normal prior.* Journal of Visual Communication and Image Representation, *3, No.4:316–324, December 1992.*

[101] C.S. Butler and M.I. Miller. *Maximum a posteriori estimation for SPECT using regularization techniques on massively-parallel computers.* IEEE Transactions on Medical Imaging, *12(1):84–89, March 1993.*

[102] Marcus E. Raichle. *Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli.* Science, *249:1041–1044, August 1990.*

[103] P.T. Fox and M.A. Mintun. *Noninvasive functional brain mapping by change-distribution analysis of averaged pet images of water tissue activity.* Clinical Sciences, *30, no. 2:141–149, February 1989.*

[104] K.J. Friston, P. Jezzard, and R. Turner. *Analysis of functional mri time-series.* Human Brain Mapping, *2:69–78, 1994.*

[105] C.A. Pelizzari, G.T.Y. Chen, D.R. Spelbring, R.R. Weichselbaum, and C.T. Chen. *Accurate three-dimensional registration of ct, pet, and/or mr images of the brain.* Journal of Computer Assisted Tomography, *13(1):20–26, 1989.*

[106] V. E. Johnson, W. H. Wong, X. Hu, and C.-T. Chen. *Image restoration using Gibbs priors: Boundary modeling, treatment of blurring, and selection of hyperparameter.* IEEE Trans. on Pattern Analysis and Machine Intelligence, *PAMI-13, No. 5:413–424, May 1991.*

[107] Y. Amit and K. M. Manbeck. *Deformable template models for emission tomography.* Repts. in Pattern Analysis, *155, 1991.*

[108] D.G. Politte and D.L. Snyder. *The use of constraints to eliminate artifacts in maximum-likelihood image estimation for emission tomography.* IEEE Transactions on Nuclear Science, *35,no.1:608–610, February 1988.*

[109] J.G. Csernansky, G.M. Murphy, and W.O. Faustman. *Lymbic/mesolimbic connections and the pathogenesis of schizophrenia.* Journal of Biological Psychology, *30:383–400, 1991.*

[110] G.E. Christensen, A.A. Kane, J.L. Marsh, and M.W. Vannier. *A 3d deformable infant ct atlas.* In Computer Assisted Radiology, in press, *Paris, June 1996.*

[111] G.E. Christensen, A.A. Kane, J.L. Marsh, and M.W. Vannier. *Synthesis of an individualized cranial atlas with dysmorhpic shape.* IEEE Proceedings of Mathematical Methods in Biomedical Image Analysis, in press, *June 1996.*

[112] D.E. Goodkin, J.S. Ross, M.S. Vanderbrug, J. Konecsni, and R.A. Rudick. *Magnetic resonance imaging lesion enlargement in multiple sclerosis. disease-related activity, chance occurance, or measurement artifact?* Arch Neurol, *49:261–263, 1992.*

[113] G.W. Roberts. *Schizophrenia: The cellular biology of a functional psychosis.* Trends Neurosci, *13:207–211, 1990.*

[114] G. Bartzokis, J. Mintz, P. Marx, D. Osborn, D. Gutkind, F. Chiang, CK Phelan, and SR Marder. *Reliability of in vivo volume measures of hippocampus and other brain structures using mri.* Magnetic Resonance Imaging, *11:993–1006, 1993.*

[115] A.J. Bartley, D. Jones, and D.R. Weinberger. *The heritability of gyral patterns in monozygotic and dizygotic twins.* Neuropsychopharmacol, *10:138S, 1994.*

[116] R. Bajcsy and S. Kovacic. *Multiresolution Elastic Matching.* Computer Vision, Graphics, and Image Processing, *46:1–21, 1989.*

[117] Y. Amit, U. Grenander, and M. Piccioni. *Structural image restoration through deformable templates.* J. American Statistical Association, *86(414):376–387, June 1991.*

[118] S. Timoshenko. Theory of Elasticity. *McGraw-Hill, 1934.*

[119] L.A. Segel. Mathematics Applied to Continuum Mechanics. *Dover Publications, New York, 1987.*

[120] G. Christensen. Deformable Shape Models for Anatomy. *Ph.D. Dissertation, Department of Electrical Engineering, Sever Institute of Technology, Washington University, St. Louis, MO, Aug 1994.*

[121] L. Younes. *Discussion of mathematics for object recognition shape, invariance and deformations.* In Workshop on Mathematics for Object Recognition Shape, Invariance and Deformations, *November 1997.*

[122] M. Claudio and S. Roberto. *Using marching cubes on small machines.* Graphical Models and Image Processing, *56:182–3, 1994.*

[123] B. Hamann. *Curvature approximation for triangulated surfaces.* In Computing, *pages 139–153. Springer-Verlag, 1993.*

[124] J.T. Kent and K. V. Mardia. *The link between kriging and thin-plate splines.* In F. P. Kelly, editor, Probability, Statistics and Optimisation. *John Wiley and Sons, 1994.*

[125] John C. Csernansky, Sarang Joshi, Lei Wang, Mokhtar Gado, J. Philip Miller, Ulf Grenander, and Michael I. Miller. *Hippocampal morphometry in schizophrenia by high dimensional brain mapping. September 1997.*

[126] U. Grenander and G. Szego. Toeplitz Forms and Their Applications. *University of California Press, Berkeley, CA., 1958.*

[127] *Hui-Hsiung Kuo.* Gaussian Measures in Banach Spaces, *volume Lecture Notes in Mathematics - 463. Springer-Verlag, 1975.*

[128] *R. Szeliski.* Bayesian Modeling of Uncertainy in Low-Level Vision. *Kluwer Academic Publisher, Boston, 1989.*

[129] *Frithiof I. Niordson.* Shell Throry, *volume 29 of* Applied Mathematics and Mechanics. *North-Holland, Amsterdam New York Oxford, 1985.*

[130] ADINA Theory and Modeling Guide. *ADINA R & D, Inc., Watertown MA, 1992.*

[131] *M. Reed and B. Simon.* Fourier Analysis, Self-Adjointness. *Academic Press, 1980.*

[132] *Y. Amit and M. Piccioni. A nonhomogenous markov process for the estimation of gaussian random fields with non-linear observations.* The Annals of Probability, *19:1664–1678, 1991.*

[133] *D.C. Van Essen and J.H.R. Maunsell. Two-dimensional maps of the cerebral cortex.* J. Comp. Neurology, *191:255–281, 1980.*

[134] *Wally Welker. Why does cerebral cortex fissure and fold.* Cerebral Cortex, *83:3–136, 1990.*

[135] *P.M. Thompson, C. Schwartz, R.T. Lin, A.A. Khan, and A.W. Toga. Three-dimensional statistical analysis of sulcal variability in the human brain.* Journal of Neuroscience, *16(13):4261–4274, 1996.*

[136] *I.R. Porteous.* Geometric Differentiation. *Cambridge University Press, Cambridge, 1994.*

[137] *D.B. Cooper, H. Elliott, F. Cohen, L. Reiss, and P. Symosek. Stochastic boundary estimation and object recognition. In A. Rosenfeld, editor,* Image Modeling, *pages 63–94. Academic Press, New York, 1981.*

[138] *J. L. Elion, S. A. Geman, and K.M. Manbeck. Computer recognition of coronary arteries.* Journal of the American College of Cardiology, *17,No.2, February 1991.*

[139] *M. Joshi.* Connected Brain Segmentation and Cortical Surface Generation. *Master's Thesis, Department of Electrical Engineering, Sever Institute of Technology, Washington University, St. Louis,Mo., August 1998.*

[140] *N. Khaneja.* Cortical Surface and Sulcal Mapping in Anatomy. *Masters Thesis, Department of Electrical Engineering, Sever Institute of Technology, Washington University, St. Louis,Mo., September 1996.*

[141] *Barrett O'Neill.* Elementary Differential Geometry. *Academic Press, Inc, San Diego, 1966.*

[142] *U. Grenander and Z. Lu. Pattern theoretic representations of anatomical pathologies. In* Technical Report. *Division of Applied Mathematics, 1994.*

[143] *U. Grenander, Y. Chow, and D. Keenan.* HANDS: A Pattern Theoretic Study of Biological Shapes. *Springer-Verlag, New York, 1990.*