Statistical Morphometry in Neuroanatomy

Moo K. Chung

Department of Mathematics and Statistics

McGill University, Montreal

May, 2001

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirement of the degree of Doctor of Philosophy

Abstract

The scientific aim of computational neuroanatomy using magnetic resonance imaging (MRI) is to quantify inter- and intra-subject morphological variabilities. A unified statistical framework for analyzing temporally varying brain morphology is presented. Based on the mathematical framework of differential geometry, the deformation of the brain is modeled and key morphological descriptors such as length, area, volume dilatation and curvature change are computed. To increase the signal-to-noise ratio, Gaussian kernel smoothing is applied to 3D images. For 2D curved cortical surface, diffusion smoothing, which generalizes Gaussian kernel smoothing, has been developed. Afterwards, statistical inference is based on the excursion probability of random fields defined on manifolds.

This method has been applied in localizing the regions of brain tissue growth and loss in a group of 28 normal children and adolescents. It is shown that children's brains change dramatically in localized areas even after age 12.

Résumé

La neuroanatomie numérique, utilisant l'imagerie par résonnance magnétique (MRI), a pour but de quantifier les variabilités morphologiques intra et inter-sujets. Une analyse statistique de la variabilité temporaire de la morphologie du cerveau est présentée. Basé sur des principes de géométrie différentielle, la déformation du cerveau est modelée, et des caractéristiques morphologiques sont calculées: les dilatations de la longueur, de la surface et du volume du cerveau, ainsi que les changements de courbure. Afin d'augmenter le ratio signal/bruit, un lissage gaussien est appliqué aux images tridimensionnelles. Pour les images des surfaces corticales en 2 dimensions, le lissage par diffusion a été utilisé; celui-ci généralise le lissage gaussien de manifolds arbitraires. L'inférence statistique est basée sur la probabilité d'excursion des champs aléatoires définis sur les manifolds.

Cette méthode a été appliquée à un groupe de 28 enfants et adolescents, afin de localiser les régions de croissance et de perte du tissus cérébral. On montrera que les cerveaux de ces enfants changent considérablement, contredisant certaines hypothéses préalables au sujet du développement du cerveau.

Acknowledgements

I would like to express my sincere gratitude to supervisors Professors Keith J. Worsley and Jim O. Ramsay, for continuous guidance and encouragement during the time of uncertainty. Without their support, I would not have made this far.

I would like to acknowledge the assistance of Steve Robbins, Tomas Paus, Alan C. Evans and Louis Collins of the McConnell Brain Imaging Center, Montreal Neurological Institute. Steve Robbins at the Montreal Neurological Institute, in particular, provided technical assistance with the cortical surface extraction and registration. A special thanks to Jay Giedd and Judith Rapoport of the National Institute of Mental Health for permission to use the original 28 subjects pediatric MRI images.

My thanks are also to Jonathan Taylor for many discussions about the finite element method and Carine Bellera for her French translation of the abstract. I am also grateful to Professors George P.H. Styan and David Wolfson for their advice and encouragement. On a practical note, I am indebted to many people at the Department of Mathematics and Statistics and the Montreal Neurological Institute over the years. Finally I would like to thank my parents for their endless love.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Deformation-Based Morphometry

Brain morphology across the span of human aging is not uniform. Brain growth spurts during childhood are followed by a brief period of morphological stability as loss of brain volume begins in middle age. The advancement of magnetic resonance imaging (MRI) gives us a new computational tool for the characterization of such temporally varying brain morphology and this is emerging as the new field of *computational neuroanatomy* [59]. MRI depends on the response of magnetic fields to produce digital images that provide structural information about brain tissue. This noninvasive but somewhat expensive procedure has become a standard neuro-imaging modality in examining the structure of the brain.

In order to characterize temporally varying brain morphology, we must compare different individual images taken at different times. Such comparisons require a reference coordinate system which can be obtained via *image registration*. The objective of the image registration problem is to deform as smoothly as possible from one brain image to another brain image. For the complete treatment of various image registration techniques, see the book "Brain Warping" edited by A.W. Toga [114] and an article by H. Lester and R.A. Simon [70]. One of the most widely used registration method is the intensity-based matching, which tries to align one image to another in such a way that the correlation of the image intensity is maximized [25, 26]. Alternate methods based on elastic deformation and fluid dynamics

models are also available [21, 32, 44, 111, 113].

Via image registration algorithms, biologically homologous points in two different images are identified and the mathematical transformation between these two points, called *deformation*, can be computed. The deformation is given as a 3-dimensional vector at each voxel. Mathematically this deformation can be represented as a transformation from a point x to a homologous point $x + U(x)$ in a fixed coordinate system. The 3-dimensional vector U is usually called *displacement field* in elastic deformation theory [75] and it measures a relative movement of the point x. Morphological studies based on studying this deformation is called the *deformation-based morphometry* [8]. Although the idea of deformation originates from elastic theory and continuum mechanics [19, 75], perhaps the first person to apply this concept to deform one biological structure to another closely related structure is D'arcy Thompson in his classical book "On Growth and Form" [109], where he deformed the skulls of human and primates, and other biological structures using deformable grids. Unlike classical morphometry in shape analysis [12, 13, 40, 62, 101] , the deformation-based morphometry tries to avoid anatomical landmarks in characterizing morphological changes. An *anatomical landmark* is a point assigned by an expert that corresponds between organism in some biologically meaningful way [40]. However, it is very hard to identify such anatomical landmarks in brain images systematically and this is the one reason why automatic image registration methods are preferred to image registration methods based on anatomical landmarks.

In many morphological studies, temporally varying morphological differences in the brain have been examined primarily by MRI-based volumetry until now. Classical MRI-based volumetry requires segmentation of the region of interest, either manually or by spatial normalization, in two MR images taken at different times t_1 and t_2. Then the total volumes V_1 and V_2 of the homologous regions are calculated by counting the total number of voxels. Afterwards, the volume variation $\Delta V = V_2 - V_1$ is used as an index of morphological changes [47, 85, 89, 107]. So the advantage of deformation-based morphometry over the classical MRI-based volumetry is that it does not require *a priori* knowledge of the region of interest to perform the morphological analysis. Moreover, the deformation-based volumetry improves the power of detecting the regions of volume change within the limits of the accuracy of the registration algorithm. These two advantages of the deformation-based volumetry over the

standard MRI-based volumetry have also been noted in [8, 32].

As a part of deformation-based morphometry, a new technique called deformation-based volumetry is emerging; this method does not require segmentation of *a priori* regions of interest [8, 32]. In deformation-based volumetry, the Jacobian of the deformation field that is required to register one brain to another is used to detect volumetric changes. By introducing the concept of local volume dilatation , which is the first order approximation of the Jacobian change, the local volume change at each voxel can be computed and used to measure possible brain tissue growth or loss [22, 23]. By definition, the Jacobian of the deformation is the volume of the unit-cube after the deformation. Assuming that one can find the deformation field at any voxel, volume change can be detected at a voxel level. Although it seems that there are many different ways of detecting morphological changes in deformation-based morphometry, a translation, a rotation and a strain are sufficient for detecting a relatively small displacement and, in turn, for characterization of morphological changes over time.

Because the deformation-based morphometry (DBM) is a relatively new method, there are very few morphological studies that have used the Jacobian for local volume-change. C. Davatzikos *et al.* used the Jacobian of the 2D deformation field as a measure of local area-change in 2D cross-sections of the corpus callosum to test gender-specific shape differences [34]. P. Thompson *et al.* applied the Jacobian of 3D deformations as a measure of the regional growth of the corpus callosum [113]. J.-P. Thirion used the divergence of the displacement-vector field, which is equivalent to the dilatation, for detecting growth of brain tumors [107]. Also P. Thompson *et al.* used local rates of dilatation, contraction and shearing from the deformation field to detect morphological changes in brain development [110]. Although these researchers are using the same or closely related concept of Jacobian, most of previous deformation-based volumetric studies lack systematic statistical treatments. The major aim of this thesis is to develop the unified statistical methodology for the deformation-based morphometry based on the mathematical framework of random fields.

1.2 Random Fields

In the deformation-based morphometry, it is sufficient to analyze the displacement vector fields U, which are assumed to be realizations of a vector of Gaussian random fields defined in either 3D Euclidean space or 2D Riemannian manifolds. The generalization of a continuous stochastic process defined in $\mathbb{R}$ to a higher dimensional abstract space is called a *random field*. For the general overview of random fields, see the books by R. J. Adler [1], E.R. Dougherty [38] and A.M. Yaglom [126]. Given a probability space, a random field $X(x)$ defined in $\mathbb{R}^n$ is a function such that for every fixed $x \in \mathbb{R}^n$, $X(x)$ is a random variable on the probability space. The more precise measure-theoretic definition can be found in [1, pp. 13]. It is also possible to extend the definition of a random field onto a Riemannian manifold [59, 106].

The *covariance function* $R(x, y)$ of a random field X is defined as the cross-covariance between two random variables $X(x)$ and $X(y)$:

$$R(x, y) = \mathbb{E}\Big((X(x) - \mu(x))(X(y) - \mu(x)) \Big),$$

where $\mu(x) = \mathbb{E}(X(x))$ is the mean function. If the joint distribution

$$F_{x^1,\cdots,x^m}(z^1, \cdots, z^m) = P\big(X(x^1) \leq z^1, \cdots, X(x^m) \leq z^m\big)$$

is invariant under the translation from $(x^1, \cdots, x^m) \to (x^1 + \tau, \cdots, x^m + \tau)$, then a random field X is said to be *stationary* [1]. For a stationary random field X, we can show $\mu(x) = \mu(0)$ and $R(x, y) = f(x - y)$ for some function f. Although the converse is not always true, such a case has never been encountered in practical applications [126]. Since we are interested in more practical applications, throughout the thesis we will equate the stationarity with the condition $\mu(x) = \mu(0)$ and $R(x, y) = f(x - y)$.

An important class of random fields is Gaussian fields. A *Gaussian random field* $X(x)$ is a random field whose finite joint distribution $F_{x^1,\cdots,x^m}(z^1, \cdots, z^m)$ is a multivariate Normal for every $x^1, \cdots, x^m$. Because any mean zero multivariate Normal distribution can be completely characterized by its covariance matrix, a mean zero Gaussian random field X can be uniquely determined by its covariance function $R(x, y)$. As in the case of Gaussian random variables, we can use Gaussian fields to construct new random fields such as χ^2, t, F and Hotelling's

T^2 fields, all of which are discussed through the thesis. For example, the χ^2 field with m degrees of freedom is defined as

$$W(x) = \sum_{i=1}^{m} X_i^2(x),$$

where $X_1, \cdots, X_m$ are independent, identically distributed Gaussian fields with zero mean and the variance equals to one [1]. Similarly, we can define t and F fields [121] as well as Hotelling's T^2 field [17]. The Hotelling's T^2 statistic for the displacement field U has been widely used in detecting morphological changes [17, 27, 43, 59, 111]. In particular, J. Cao and K.J. Worsley in [17] were able to derive the excursion probability of the Hotelling's T^2 random field and applied it to detect gender specific morphological differences. However, the Hotelling's T^2 field only measures the amount of relative translation and it is not sufficient to characterize local morphological changes.

Statistical inferences for a random field X have been usually based on the series expansion of X of the form

$$X(x) = \sum_{i=1}^{\infty} Z_i \phi_i(x),$$

where $\{\phi_i(x)\}$ are basis functions and $\{Z_i\}$ are random variables. Then the statistical inference is performed on the realizations of coefficients $\{Z_i\}$ [11]. For a Gaussian random field, the most well known series expansion is called the *Karhunen-Loeve expansion* [38, 66, 126]. Alternately, statistical inference can be based on the extreme distributions of X [1, 69]. For a finite domain $\Omega \in \mathbb{R}^N$, a random variable $\max_{x \in \Omega} X(x)$ has been used as a test statistic in both functional and structural brain imaging studies [4, 17, 23, 119, 125].

1.3 Cortical Surface

We can extend the deformation-based morphometry to the cortical surfaces. Brain tissue is usually classified into two types: *gray matter* (GM) and *white matter* (WM) (Figure 1.1). The empty space between the brain and the skull is filled with a liquid called *cerebral spinal fluid* (CSF). The *cerebral cortex* (outer cortical surface) is the boundary layer between CSF and gray matter. It has the topology of a 2-dimensional highly convoluted sheet and the most of the cortical surface is buried deep inside. This has been partially confirmed by computing the total surface area of the outer cortical surface in the thesis. Most of the features that

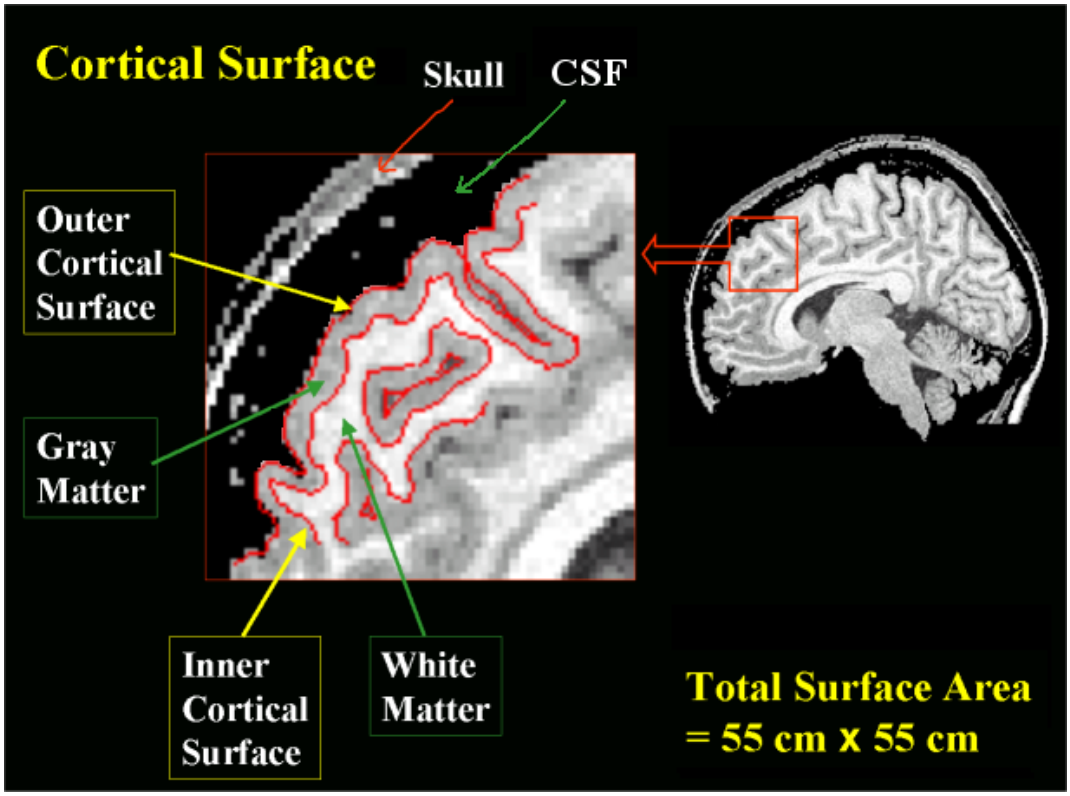

Figure 1.1: Cortical anatomy.

distinguish these cortical regions can only be measured relative to the local orientation of the cortical surface [31]. The inner cortical surface is the boundary layer between gray matter and white matter. As brain develops over time, cortical surface area as well as cortical thickness and the curvature of the cortical surface change. It is highly likely that such age-related changes of the cortical surface area and cortical thickness are not uniform (Figure 1.2). By measuring how such morphological descriptors change over time, brain tissue growth or loss of cortical regions can be localized [24]. The first obstacle in developing surfaced-based morphometry is automatic *segmentation* or extraction of the cortical surfaces from MRI . The segmented surface is usually represented as a triangular mesh. The most widely used method for triangulating the surface is the marching cubes algorithm [72]. Level set method [97] or deformable surfaces method [33] are also available. In our thesis, we have used the anatomic segmentation using proximities (ASP) method [73, 74], which is a variant of the deformable surfaces method, in order to generate cortical triangular meshes consisting of 81,920 triangles. Once we have a triangular mesh as the realization of the cortical surface, we can model the cortical surface deformation. In modeling the surface deformation, a proper mathematical framework can be found in both differential geometry and fluid dynamics. The

concept of the *evolution of phase-boundary* in fluid dynamics [39, 53], which describes the geometric properties of the evolution of boundary layer between two different materials due to internal growth or external force, can be used to derive the mathematical formula for surface deformation.

It is natural to assume the cortical surfaces to be a smooth 2-dimensional *Riemannian manifold* parameterized by u^1 and u^2:

$$X(u^1, u^2) = \{x_1(u^1, u^2), x_2(u^1, u^2), x_3(u^1, u^2) : (u^1, u^2) \in D \subset \mathbb{R}^2\}.$$

A more precise definition of a Riemannian manifold and a *parameterized surface* can be found in the classical books by W.M. Boothby [14], M.P.D. Carmo [18] and E. Kreyszig [64]. The pair of the mapping X and its range $X(D)$ is called a coordinate neighborhood or a *local chart*. A local chart can be approximated at each point on the surface by locally approximating the diffeomorphism by up to the quadratic terms in its Taylor series. If D is a unit square in $\mathbb{R}^2$ and a surface is topologically equivalent to a sphere then at least two different local charts are required. Due perhaps to the computational difficulty of extracting the cortical surfaces of human brain, there are not many studies that models the cortical surfaces as Riemannian manifolds. Gaussian and mean curvatures of the brain surface have been used to characterize its shape [33, 50, 60]. In particular, S.C. Joshi *et al.* used the quadratic surface in estimating the Gaussian and mean curvature of the cortical surfaces [60]. S. Angenent *et al.* used a conformal mapping to flatten the brain surface in a way which preserves angles [6]. Bakircioglu *et al.* and P. Thompson *et al.* used spherical harmonics in cortical surface registration [10, 112].

1.4 Image Smoothing

All brain images are inherently noisy due to errors associated with image acquisition. Compounding the image acquisition errors, there are errors caused by image registration and surface parameterization. In order to increase the SNR (signal-to-noise ratio), image smoothing is most often used. The SNR is defined as the ratio:

$$\text{SNR} = \frac{\text{Variance of signal}}{\text{Variance of noise}}.$$

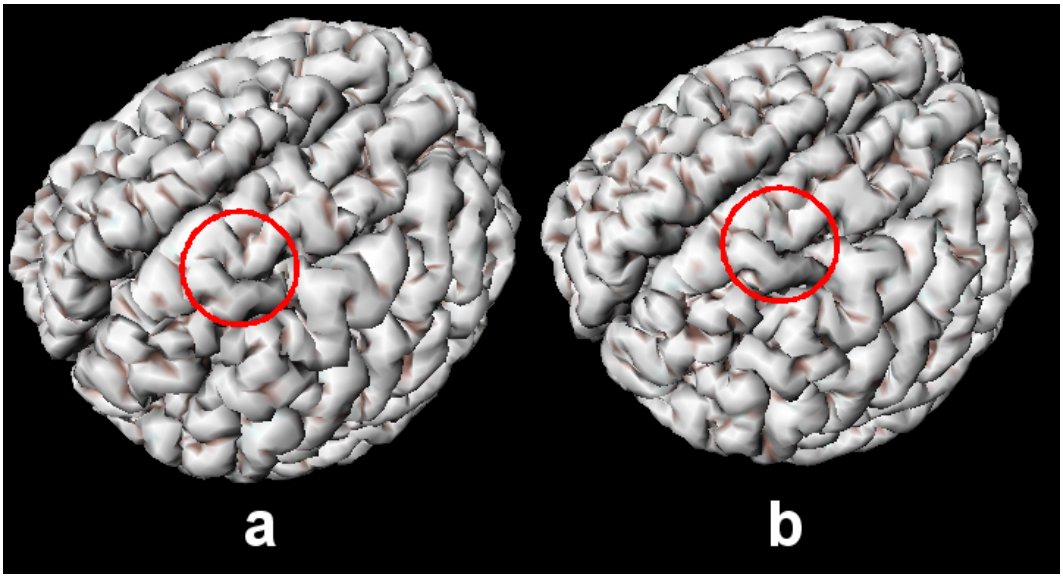

Figure 1.2: Outer cortical surface of a single subject **a.** at age 14 and **b.** at age 19 showing globally similar cortical patterns although slight sulcal variations can be found inside the marked circles. On the other hand, cortical surfaces of different individuals show more sulcal variabilities.

The precise definition of SNR using the spectral density can be found in [38, 91, 124]. Among many possible image smoothing methods, *Gaussian kernel smoothing* has emerged as a de facto smoothing technique among brain imaging researchers [63, 79]. An integral version of Gaussian kernel smoothing of an n-dimensional image $f(\mathbf{x}), \mathbf{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$ with smoothing parameter $h > 0$ is defined by

$$F^*(\mathbf{x}, h) = \int_{\mathbb{R}^n} K\Big(\frac{\mathbf{x} - \mathbf{y}}{h}\Big) \frac{f(\mathbf{y})}{h^n}\, d\mathbf{y},$$

where a Gaussian kernel is $K(\mathbf{x}) = (2\pi)^{-n/2} \exp(-\|\mathbf{x}\|^2/2)$. $F^*(\mathbf{x}, h)$ is the *scale-space* representation of the image $f(\mathbf{x})$ first introduced by A. Witkin [117]. Gaussian kernel smoothing, as the name implies, tends to blur images as h gets large. So each $F^*(\mathbf{x}, h)$ for different values of h produces a blurred copy of its original. The resulting scale-space representation from coarse to fine resolution can be used in multiscale image processing approaches such as hierarchical searches and image segmentation. See articles [71, 81, 82, 83, 98, 124] for the review of the major problems in scale-space and multiscale descriptions of images. $F(\mathbf{x}, t) = F^*(\mathbf{x}, \sqrt{2t}\,)$ is known to satisfy the diffusion equation

$$\frac{\partial F}{\partial t} = \Delta F, \;\; F(\mathbf{x}, 0) = f(\mathbf{x}),$$

where $\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$ is the Laplacian in $\mathbb{R}^n$. See an article by A. Grigor'yan [51] for a brief overview of the relations between a Gaussian kernel, diffusion equations and Brownian

motion. Hence, Gaussian kernel smoothing can be derived by solving a diffusion equation. The most widely used methods in solving diffusion equations are the *finite element method* (FEM) and the *finite difference method* (FDM). The earliest mathematical treatment of the FEM can be found in an article by R. Courant [28]. It has been a standard tool for solving PDEs and variational problems. The books by M.N.O. Sadiku [93], and J. Oden and G. Garey [77] would serve as good text books for the FEM. Also see the book by J.C. Strikwerda for the review of FDM [102]. The FEM usually requires four steps [92]: 1) discretizing the solution region into finite number of subregions called *finite elements*, 2) deriving governing equations for an element, 3) assembling of all elements, 4) and solving the system of equations.

The drawback of Gaussian kernel smoothing is that it does not respect the natural boundaries of objects. We would like to encourage smoothing within a region rather than smoothing across the boundaries. This could be achieved by solving the diffusion equation with the condition $F(\mathbf{x}, t) = 0$ on the boundaries. Solving a partial differential equation with such boundary condition is called the BVP (boundary value problem) and such smoothing method is usually referred as *diffusion smoothing* or diffusion filtering. J.O. Ramsay has solved the BVP to smooth data constrained within a region [86]. Extending the work of A. Witkin [117], P. Pernona and J. Malik [79] first introduced the concept of anisotropic diffusion in the problem of edge enhancement and detection by running the diffusion equation backwards in time. Diffusion smoothing has been also used in the analysis of functional magnetic resonance imaging (fMRI) data on the brain surface [4] and detecting the regions of surface area change in brain development [24].

1.5 Outline of the Thesis

Chapter 2 presents the deformation-based morphometry in 3D whole brain MRI and introduce the concept of local volume dilatation, which will be modeled in the mathematical setting of random fields. The statistical methodology will be tested to normal brain development and comparisons will be made with previous brain developmental studies.

Chapter 3 contains the further generalization of diffusion smoothing in $\mathbb{R}^n$ to an arbitrary

Riemannian manifold using the Laplace-Beltrami operator. It is essential to have surface-based smoothing algorithm that generalizes Gaussian kernel smoothing, in order to apply the results of random field theory without major modification. We will present two different methods to solve a diffusion equation on manifolds. The first method uses quadratic polynomials for local surface parameterization. The second method will be based on the finite element method, which will be used to derive the exact mathematical form of the estimated Laplace-Beltrami operator.

Chapter 4 extend the deformation-based morphometry developed in Chapter 2 to the cortical surfaces by modeling the cortical surface as a Riemannian manifold. Morphological descriptors based on the geometry of the manifold will be used to detect surface area, cortical thickness and curvature changes over time.

Finally, Chapter 4 presents a brief summary of the thesis as well as the future research topics which have been left out in the thesis.

Chapter 2

Deformation-Based Morphometry in 3D Volumes

In this chapter, we present a unified statistical framework for detecting brain tissue growth and loss in temporally varying brain morphology. As an illustration, we will demonstrate how the method can be applied in detecting regions of tissue growth and loss in brain images longitudinally collected in a group of the same children and adolescents.

2.1 Deformation Model

Unlike other brain morphological studies that try to characterize the structural variabilities among different individuals of similar age groups, morphological studies of temporally varying brain structure have an extra temporal dimension. Therefore, a different approach to morphometry is required to fully understand the spatio-temporal complexity of brain development.

Let $U(x,t) = (U_1, U_2, U_3)$ be the 3D displacement vector field required to move the structure at position $x = (x_1, x_2, x_3)$ of the atlas or template brain Ω_{atlas} and at the reference time 0 to the corresponding position after time t. Thus the structure at x deforms to $x + U(x,t)$ with respect to a fixed reference coordinate. The displacement U at fixed time t is usually estimated via volume-based non-linear registration techniques on two images

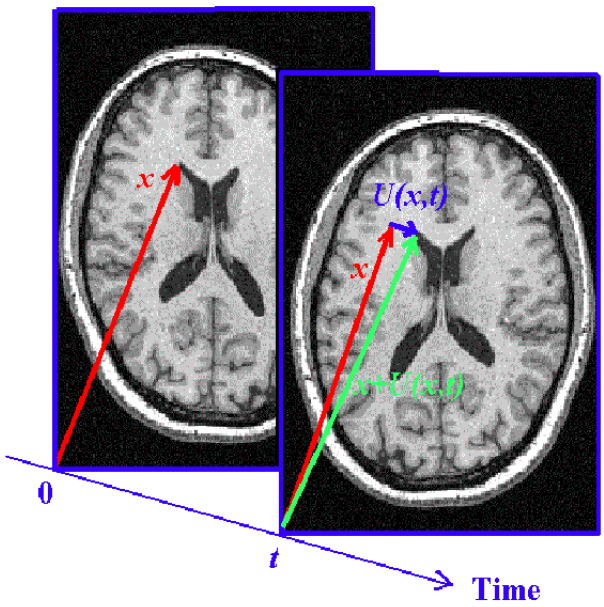

Figure 2.1: The structure at x deforms to the homologous structure at $x+U(x,t)$ after time t.

taken at time 0 and t. Then we propose to test the following stochastic model of brain development:

$$\frac{\partial U}{\partial t}(x,t) = L(U) + \Sigma^{1/2}(x)\epsilon(x),\ x \in \Omega_{atlas}, t \in \mathbb{R}^+ \tag{2.1}$$

where L is a partial differential operator involving spatial components and $\Sigma(x)$ is the 3 $\times$ 3 symmetric positive-definite covariance matrix, which allows for correlations between components of the deformations and depends on the spatial coordinates x only. Since Σ is symmetric positive-definite, the square-root of Σ always exists. The components of the error vector ϵ are independent and identically distributed as smooth stationary Gaussian random fields with zero mean and unit standard deviation. The error structure $\Sigma^{1/2}\epsilon$ was first introduced in [122] and [17]. An equation of the type (2.1) is usually called a *stochastic evolution equation* and it models how the structure evolves over time. Any smooth morphological change can be completely described with the stochastic evolution equation (2.1) within the bound set by the error structure $\Sigma^{1/2}\epsilon$. If the deformation is assumed to follow a diffusing behavior, then L can be chosen as the Laplacian $L = \sigma^2(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\frac{\partial^2}{\partial x_3^2})$. If the morphological changes are assumed to follow a fluid dynamics model, L becomes a Navier-Stokes operator given in [67]. Since the displacement U is a function of both time t and space x, the partial differential operator L applied to U is again a function of both time and space.

Longitudinal analysis based on (2.1) is essentially the inverse problem of brain registration. This analysis tries to determine the partial differential operator L from given displacement fields. On the other hand, in brain registration, the objective is to find the displacement field U that matches homologous points between two images based on minimizing a cost function or actually solving partial differential equations. The most widely used physical models that have been used in brain registration are elastic deformations and fluid dynamics models [21, 32, 44, 113]. Suppose that the displacement field U is obtained as a solution of the elastic deformation equation given by

$$\frac{\partial U}{\partial t} = L_{elastic}(U) + \Sigma^{1/2}\epsilon,$$

where the operator

$$L_{elastic}(U) = \lambda_1 \nabla^2 U + \lambda_2 \nabla(\nabla \cdot U) + F$$

and ∇ is the gradient operator, $\nabla\cdot$ is the divergence operator and λ_1,λ_2 are called the Lamé constants [116]. Then using this displacement field U as given data, we try to estimate (2.1) which minimizes a certain error criterion based on $\Sigma^{1/2}\epsilon$. Then the best estimator of L is heavily biased toward the prior operator $L_{elastic}$. It indicates that the estimation of the evolution equation should be based on an image registration method that does not assume an *a priori* physical model or on an empirical Bayesian framework. We will use intensity-based registration algorithms that do not have explicit physical model assumptions to warp one brain to another [26, 9], but there should be further comparative studies of the different image registration methods to draw any general conclusions.

It can be assumed that, in the case of morphological changes occurring in a healthy brain over a relatively short period of time, deformation occurs continuously and smoothly, so the higher order temporal derivatives of the displacement U are relatively small compared to the displacement itself. In such a case, the first order approximation to $L(U)$ is sufficient to capture most of the morphological variabilities over time. Therefore, we approximate $L(U)$ with only a first order term $\mu_0(x)$, which is constant over time, i.e.

$$\frac{\partial U}{\partial t}(x,t) = \mu_0(x) + \Sigma^{1/2}(x)\epsilon(x). \tag{2.2}$$

By taking the expectation $\mathbb{E}$ on the both sides of (2.2), we see that $\mu_0 = \mathbb{E}\big(\frac{\partial U}{\partial t}\big)$, the mean displacement rate. Under the linear model (2.2), the problem of detecting local displacement

can be solved with a simple hypothesis testing problem:

$$H_0: \quad \mu_0(x) = 0 \text{ for all } x \in \Omega_{atlas}$$

vs.

$$H_1: \quad \mu_0(x) \neq 0 \text{ for some } x \in \Omega_{atlas}.$$

If one wishes to assess the convexity of the growth curve, we may use a second-order model:

$$\frac{\partial U}{\partial t} = \mu_0(x) + \mu_1(x)t + \Sigma^{1/2}(x)\epsilon(x).$$

Unlike estimating the first-order linear term μ_0, the problem of estimating the second-order nonlinear term μ_1 requires a large amount of data to have a statistically stable result due to intra-subject variabilities across spatial and temporal dimensions. We have limited our discussion to the detection of the first order morphological changes and we will not attempt to analyze the full model (2.1) in our thesis.

2.2 Detecting Local Displacement

We are interested in detecting regions of statistically significant changes in displacement using the linear model (2.2). This is a standard multivariate statistical inference problem and can be solved using the *Hotelling's T^2 statistic* [17, 43, 59, 111].

Let $U^j(x, t_j)$ be the 3D displacement vector field required to deform the structure $x \in \Omega_0^j$ to the corresponding homologous position after time t_j, where Ω_0^j is the whole brain volume of subject j. Let $V^j = \frac{U^j(x,t_j)}{t_j}$ be the displacement velocity. Then the sample mean displacement velocity $\bar{V}$ is given by

$$\bar{V}(x) = \frac{1}{n}\sum_{j=1}^{n} V^j(x)$$

while the sample covariance matrix C of the displacement velocity is given by

$$C(x) = \frac{1}{n-1}\sum_{j=1}^{n}\left(V^j(x) - \bar{V}(x)\right)\left(V^j(x) - \bar{V}(x)\right)^t,$$

where the superscript ^{t} denotes the matrix transpose. Then the Hotelling's T^2 field $H(x)$ is defined as

$$H(x) = n\bar{V}^t(x)C^{-1}(x)\bar{V}(x). \tag{2.3}$$

At each voxel x, under the hypothesis of no mean displacement velocity, i.e. $\mu_0(x) = 0$, $H(x)$ is distributed as a multiple of an F-distribution with 3 and $n-3$ degrees of freedom, i.e.

$$H(x) \overset{D}{\sim} 3\frac{n-1}{n-3}F_{3,n-3}.$$

Then the p-value of the maxima of $H(x)$, which corrects for searching across a whole brain volume, is used to localize the region of statistically significant structural displacement [17].

As pointed out in [8], the Hotelling's T^2 statistic based on the displacement does not directly localize regions within different structures, but rather identifies brain structures that have translated to different positions. It measures relative position of two particular voxels before and after the deformation. Therefore, in the context of temporally varying brain morphology, where the brain tissue growth is an important concern, the statistic based on the displacement field should be taken as an indirect measure of brain growth. The more direct morphological criterion that corresponds to the actual brain tissue growth is the Jacobian of the deformation field, which we will look at in the next section.

2.3 Detecting Local Volume Change

The deformation in the Lagrangian coordinate system i.e. fixed coordinate system at time t is

$$x \to x + U(x,t).$$

The local volume-change of the deformation in the neighborhood $x \in \Omega_t$ and at time t is determined by the Jacobian J which is defined as $J(x,t) = \det(I + \nabla U)$, where I denotes an identity matrix and ∇U is the 3×3 *displacement gradient* matrix of U given by

$$\nabla U = \begin{pmatrix} \frac{\partial U_1}{\partial x_1} & \frac{\partial U_1}{\partial x_2} & \frac{\partial U_1}{\partial x_3} \\ \frac{\partial U_2}{\partial x_1} & \frac{\partial U_2}{\partial x_2} & \frac{\partial U_2}{\partial x_3} \\ \frac{\partial U_3}{\partial x_1} & \frac{\partial U_3}{\partial x_2} & \frac{\partial U_3}{\partial x_3} \end{pmatrix}. \tag{2.4}$$

The component $\frac{\partial U_j}{\partial x_i}$ is called the *displacement tensor* and, in tensor-based morphometry [8], these nine components form scalar fields used to measure the second-order morphological variabilities. Note that local translation captures the first-order morphological variability. [8] separated the deformation-based morphometry into that utilizing the displacement fields

and that utilizing the displacement tensor. However, in our statistical framework, we have only one statistical model on the displacement and a model for the displacement gradient ∇U can be directly derived from (2.1) by taking the partial derivative with respect to the spatial coordinates x. Hence, by modeling the morphological changes in the mathematical framework of random fields [1], the situation of having two possibly incompatible statistical models on the displacement U and the displacement gradient ∇U can be avoided. In our unified statistical modeling approach using (2.1), all possible statistical distributions of morphological test criteria can be directly derived and easily manipulated from (2.1).

Since the Jacobian J measures the volume of the unit-cube after deformation, the rate of the Jacobian change, i.e. $\frac{\partial J}{\partial t}$ is the rate of the local volume change. In brain imaging, a voxel can be considered as the unit-cube; therefore, $\frac{\partial J}{\partial t}(x)$ essentially measures the amount of change in the volume of voxel x during the deformation. Expanding the Jacobian J, we get

$$\begin{aligned} J &= \det(I + \nabla U) \\ &= 1 + \mathrm{tr}(\nabla U) + \mathrm{detr}_2(\nabla U) + \det(\nabla U), \end{aligned}$$

where $\mathrm{detr}_2(\nabla U)$ is the sum of 2×2 principal minors of ∇U [54]. For relatively small displacements, which is the case in brain development, we may neglect the higher order terms and get $J \approx 1 + \mathrm{tr}(\nabla U)$. Taking the partial derivative with respect to the temporal coordinate t, we get

$$\begin{aligned} \frac{\partial J}{\partial t} &\approx \frac{\partial^2 U_1}{\partial t \partial x_1} + \frac{\partial^2 U_2}{\partial t \partial x_2} + \frac{\partial^2 U_3}{\partial t \partial x_3} \\ &= \frac{\partial}{\partial t}(\nabla \cdot U) \\ &= \nabla \cdot \Big(\frac{\partial U}{\partial t}\Big), \end{aligned}$$

where $\nabla\cdot$ is the divergence operator. In elastic theory, the *volume dilatation* is defined as $\Theta_{volume}(x) = \nabla \cdot U$ [75]. Therefore, the rate of the Jacobian change is approximately the rate of the volume dilatation change for relatively small displacements, i.e.

$$\frac{\partial J}{\partial t} \approx \frac{\partial}{\partial t}\Theta_{volume}(x) = \Lambda_{volume}(x), \tag{2.5}$$

where we term Λ_{volume} to be the volume dilatation rate. Since derivatives of a Gaussian field and the sum of components of a multivariate Gaussian field are again Gaussian field, from (2.2), we have a linear model on the volume dilatation rate Λ_{volume} given by

$$\Lambda_{volume}(x) = \lambda_{volume}(x) + \epsilon_{volume}(x), \tag{2.6}$$

where λ_{volume} is the mean volume dilatation rate and ϵ_{volume} is a Gaussian random field with zero mean. When $\lambda_{volume}(x) = 0$ in the neighborhood of x, the deformation is *incompressible* so there is no volume change. However, if $\lambda_{volume}(x) > 0$, the volume increases while $\lambda_{volume}(x) < 0$, the volume decreases after the deformation. In certain registration algorithms, the Jacobian J is forced to be larger than a certain threshold to ensure the homologous correspondence between two brains [20]. When such a registration algorithm is used, the power of detecting the region of statistically significant volume change may be reduced. Statistical inference on the linear model (2.6) is easier than that of (2.2) since it is a univariate Gaussian.

Let Θ^j_{volume} denote the volume dilatation of the displacement $U^j = (U^j_1, U^j_2, U^j_3)$ for subject j after time t_j. Then the volume dilatation rate or growth rate Λ^j_{volume} of subject j is

$$\Lambda^j_{volume} = \frac{1}{t_j}\Theta^j_{volume} = \frac{1}{t_j}\Big(\frac{\partial U^j_1}{\partial x_1} + \frac{\partial U^j_2}{\partial x_2} + \frac{\partial U^j_3}{\partial x_3}\Big).$$

In the actual numerical implementation, the displacement tensor $\frac{\partial U^j_i}{\partial x_i}$ can be computed by the finite difference on rectangular grid. For example, at voxel position $x = (x_1, x_2, x_3)$,

$$\frac{\partial U^j_1}{\partial x_1} \approx \frac{U^j_1(x_1 + \delta x_1, x_2, x_3) - U^j_1(x_1, x_2, x_3)}{\delta x_1},$$

where δx_1 is the length of the edge of a voxel along the x_1 axis. Then the T random field is defined as

$$T(x) = \sqrt{n}\frac{M_{volume}(x)}{S_{volume}(x)}, \tag{2.7}$$

where M_{volume} and S_{volume} are the sample mean and standard deviation of Λ^j_{volume}. Under the assumption of no local volume change at x, i.e. $\lambda_{volume}(x) = 0$, $T(x) \overset{D}{\sim} t_{n-1}$, a student t-distribution with $n-1$ degrees of freedom. As we shall see in the Result section, the sample mean dilatation-rate $M_{volume}(x)$ does not provide accurate information about where the brain growth is dominant but the $T(x)$ does (Figure 2.2). Then the p-value of the

maxima of $T(x)$, which corrects for searching across a whole brain volume, can be used to localize the region of statistically significant structural displacement [121]. In order to use the p-value based on the maxima of T random field, we need a certain assumption on the temporal correlation structure of the deformation field U. It is required that the variance $\mathbf{Var}(\epsilon_{volume})$ in (2.6) does not depend on time t. Strictly speaking the error ϵ_{volume} must be dependent on both space and time, but with the condition that $\epsilon_{volume} = \epsilon_{volume}(x,t)$ is stationary in time, we can suppress the dependency on time, i.e.

$$\mathbf{Var}\big(\epsilon_{volume}(x,t)\big) = \mathbf{Var}\big(\epsilon_{volume}(x,0)\big).$$

2.4 Statistical Inference in 3D whole brain volume

Statistical inference on both the T random field and the Hotelling's T^2 random field is based on the excursion probability

$$P\Big(\max_{x\in\Omega} Z(x) > z\Big) \tag{2.8}$$

where Z is a sufficiently smooth isotropic random field in $\mathbb{R}^N$. The smoothness of a random field corresponds to the random field being differentiable. According to R. J. Adler, there are very few cases for which exact formulas for the excursion probability (2.8) is known [2]. For this reason, approximating the excursion probability is essential. From the *Poisson clumping heuristic* [3, pp. 1-6],

$$P\Big(\max_{x\in\Omega} Z(x) < z\Big) \approx \exp\bigg(-\frac{\|\Omega\|}{\mathbb{E}\|A_z\|} P\big(Z(x) \geq z\big)\bigg),$$

where $\|\cdot\|$ is the Lebesgue measure of a set and the random set $A_z = \{x : Z(x) > z\}$ is called the *excursion set* above the threshold z. This approximation involves unknown $\mathbb{E}\|A_z\|$, which is the mean clump size of the excursion set. The distribution of $\|A_z\|$ has been estimated for the case of Gaussian [3], χ^2, t and F fields [16] but for general random fields, no approximation is available. An alternate approximation based on the expected Euler characteristic (EC) of A_z is also available. For very high threshold z, it can be shown that

$$P\Big(\max_{x\in\Omega} Z(x) > z\Big) \approx \mathbb{E}\ \chi(A_z),$$

where $\chi(A_z)$ is the Euler characteristic of A_z [125]. Compared to other approximation methods such as the Poisson clump heuristic and the tube formulae, the advantage of using the Euler characteristic formulation is that a simple exact expression can be found for $\mathbb{E}\ \chi(A_z)$. If Z is isotropic,

$$\mathbb{E}\ \chi(A_z) = \sum_{i=0}^{N} \phi_i(\Omega)\rho_i(z), \tag{2.9}$$

where ρ_i is the i-dimensional *EC density* of the random field Z and $\phi_i(\Omega)$ is the *Minkowski functional* of Ω [119]. Let $K_{\partial\Omega}$ be the curvature matrix of $\partial\Omega$ and $\mathrm{detr}_i(K_{\partial\Omega})$ be the sum of the determinant of all $i \times i$ principal minors of $K_{\partial\Omega}$. For $i = 0, \cdots, N-1$ the Minkowski functional $\phi_i(\Omega)$ is defined as

$$\phi_i(\Omega) = \frac{\Gamma(\frac{N-i}{2})}{2\pi^{\frac{N-i}{2}}} \int_{\partial\Omega} \mathrm{detr}_{N-1-i}(K_{\partial\Omega})\ dA,$$

and $\phi_N(\Omega) = \|\Omega\|$, the Lebesgue measure of Ω. Also the EC density ρ_i is defined as

$$\rho_i(z) = \mathbb{E}\big[(Z \geq z)\det(-\ddot{Z}_i)|\dot{Z}_i = 0\big] f_{\dot{Z}_i}(0),$$

where $\dot{Z}$ and $\ddot{Z}$ are the first and second derivatives of Z and the subscript i represents the first i components of Z and $f_{\dot{Z}_i}$ is the multivariate density of $\dot{Z}_i$. For a Gaussian field with zero mean and unit variance,

$$\rho_i(z) = \begin{cases} \int_z^\infty (2\pi)^{-\frac{1}{2}} e^{-u^2/2}\ du & , i = 0 \\ c^{\frac{i}{2}} (2\pi)^{-\frac{i+1}{2}} \mathbf{He}_{i-1}(z) e^{-z^2/2} & , i \geq 1 \end{cases},$$

where $c = \mathbf{Var}(\dot{Z}_1)$ and $\mathbf{He}_i$ is the Hermite polynomial of degree i. The exact expression for the EC density ρ_i can be found for other random fields such as t, χ^2, F fields [121], Hotelling's T^2 fields [17] and scale-space random fields [98]. In each case, the EC density ρ_i is proportional to $c^{\frac{i}{2}}$ and it changes depending on the smoothness of the field. If the random field Z is given as the convolution of a smooth kernel $K_h(x) = K(x/h)/h^N$ with a white Gaussian noise [98, 123], the covariance matrix of $\dot{Z} = \frac{\partial Z}{\partial x}$ is given by

$$\mathbf{Var}(\dot{Z}) = \frac{\int_{\mathbb{R}^N} \dot{K}(\frac{x}{h})\dot{K}^t(\frac{x}{h})\ dx}{h^2 \int_{\mathbb{R}^N} K^2(\frac{x}{h})\ dx}.$$

Applying it to a Gaussian kernel $K(x) = (2\pi)^{-n/2} e^{-\|x\|^2/2}$ gives $c = \mathbf{Var}(\dot{Z}_1) = \frac{1}{2h^2}$. In terms of the full width at half maximum (FWHM) of the kernel K_h as defined in (3.2), $c = \frac{4 \ln 2}{\mathrm{FWHM}^2}$.

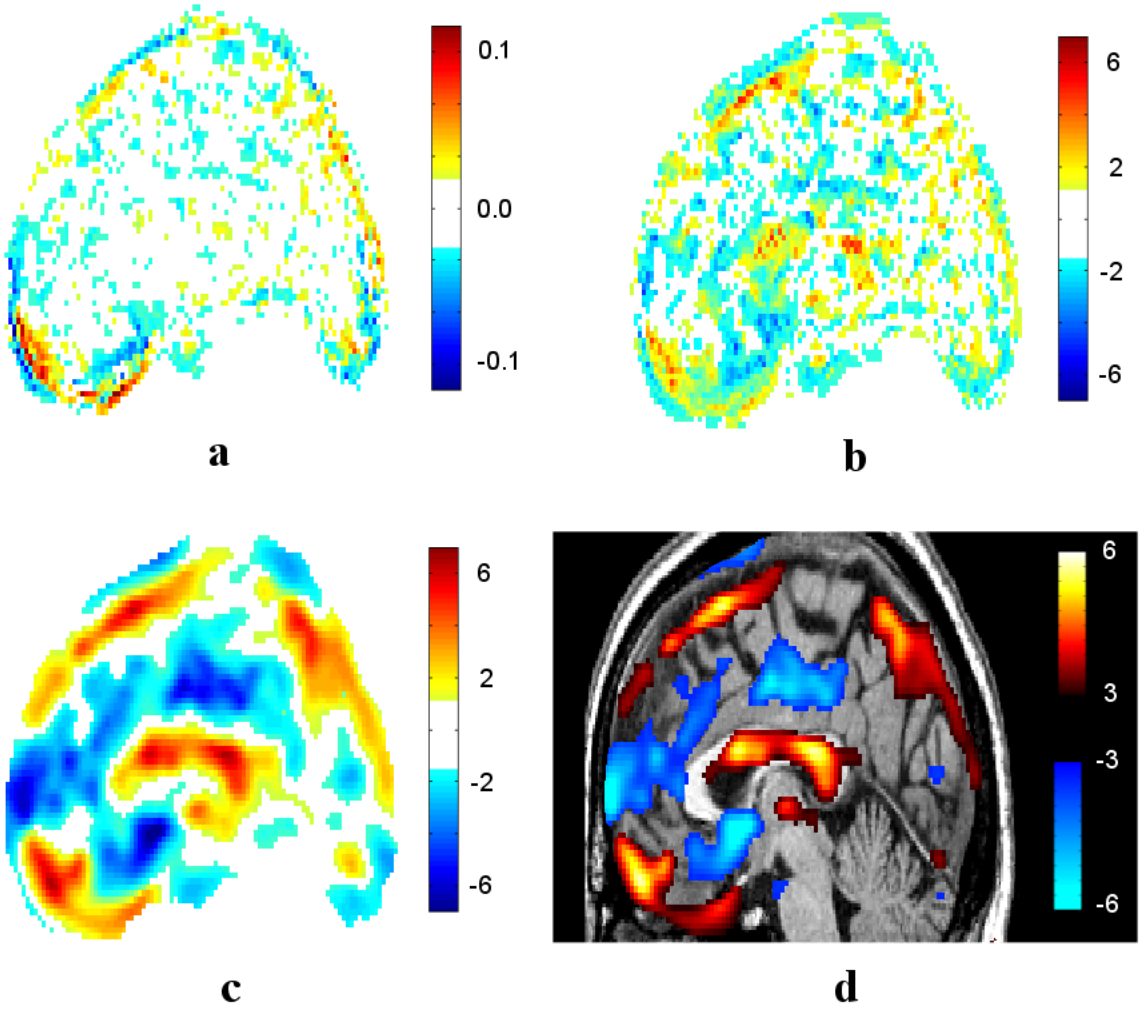

Figure 2.2: The statistical analysis of local volume change data on the mid-sagittal section. **a.** The sample mean dilatation-rate M_{volume}. It gives an incorrect impression that the local volume change only occurs near the outer cortical boundaries due, perhaps, to registration error. **b.** t-map of local volume change. Local maxima appear around the corpus callosum. A lot of noise on the cortical boundaries disappears. **c.** t-map of local volume change after 10mm Gaussian kernel smoothing. The smoothing is applied directly to the displacement fields and the signal-to-noise ratio improves. **d.** Thresholded t-map superimposed on the mid-sagittal section of the atlas brain. The corpus callosum shows volume increase. When the corrected threshold of $t > 6.5$ is applied, most of the red regions disappears except the local maximum in the splenium and the isthmus of the corpus callosum.

2.5 Important Measures in Brain Development

We have presented two different statistics (2.3) and (2.7), based on local translation and local volume changes to measure morphological changes over time. One might ask if these two statistics are sufficient to capture temporally varying morphological changes in brain and how one statistic is related to the other. Do they measure common morphological properties or different aspects of morphological changes? In this section, we will give some answers to these questions.

For a relatively small displacement, neglecting higher order terms in the Taylor expansion, the displacement U at $x + dx$ can be written as

$$U(x + dx, t) \approx U(x, t) + \nabla U(x, t)\, dx.$$

As we have pointed out already, some elements of ∇U are used to measure morphological changes [107, 8, 110]. The displacement tensor can be further decomposed into two parts depending on whether it is symmetric or antisymmetric:

$$\frac{\partial U_j}{\partial x_i} = \frac{1}{2}\Big(\frac{\partial U_j}{\partial x_i} - \frac{\partial U_i}{\partial x_j}\Big) + \frac{1}{2}\Big(\frac{\partial U_j}{\partial x_i} + \frac{\partial U_i}{\partial x_j}\Big). \tag{2.10}$$

The antisymmetric first part corresponds to a rotation or *vorticity* [94] of the deformation and the symmetric second part corresponds to a *strain*. Then the displacement U can be decomposed into three parts:

$$U(x + dx, t) \approx U(x, t) - \omega(x, t) \times dx + \varepsilon(x, t)\, dx,$$

where $\omega = \frac{1}{2}(\nabla \times U)$ is the vorticity vector and $\varepsilon = (\varepsilon_{ij}) = \frac{1}{2}\big(\nabla U + (\nabla U)^t\big)$ is the strain matrix. By taking the temporal derivative, we have the displacement velocity V decomposed into three parts:

$$V(x + dx, t) \;\approx\; V(x, t) - \frac{\partial w}{\partial t}(x, t) \times dx + \frac{\partial \varepsilon}{\partial t}(x, t)\, dx. \tag{2.11}$$

(2.11) captures most of the spatio-temporal variabilities of the displacement velocity into three components: the rate of changes in a translation, a rotation and a strain for relatively small displacements.

The strain-rate tensor $\frac{\partial \varepsilon_{ij}}{\partial t}$ can be further separated into two parts: the diagonal elements $\frac{\partial \varepsilon_{ii}}{\partial t}$ describing the length change of the volume element in each x_1, x_2 and x_3 coordinate, and

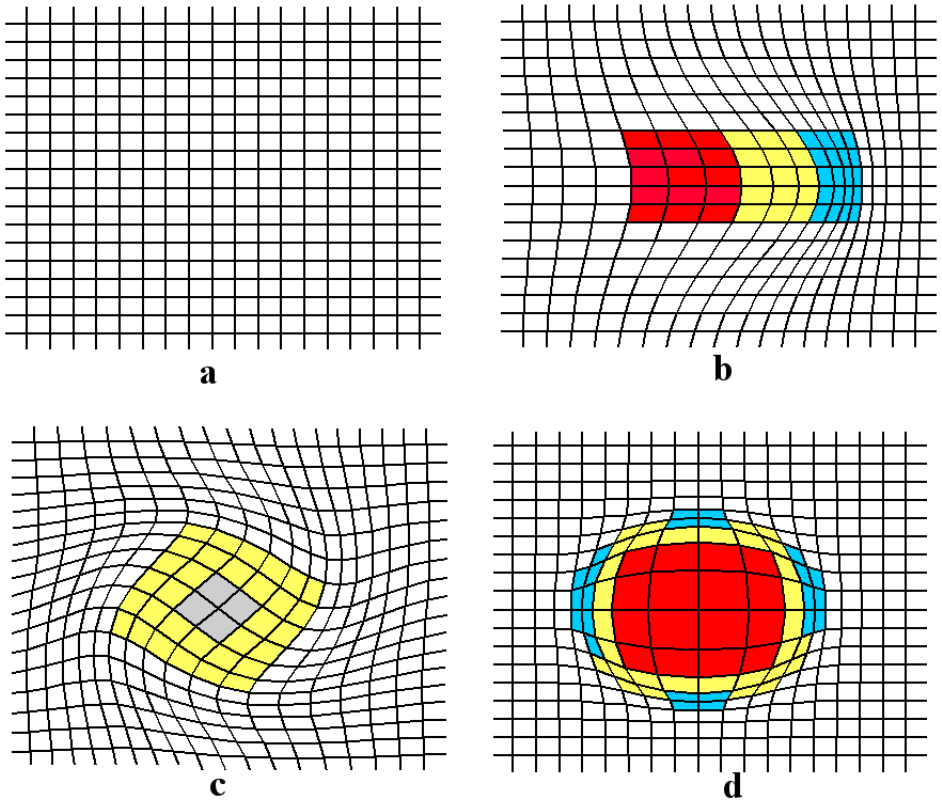

Figure 2.3: Square grid under translation, rotation and volume change. Red: volume increase, Blue: volume decrease, Gray: rotation, Yellow: translation. **a.** square grid under no deformation. **b.** Horizontal translation caused by local volume increase on the left side. **c.** 45 degree clockwise rotation. The rotation induces the outer region of the center of the rotation to translate. **d.** Volume expansion in the middle causes the grid to radially translate outward.

the off-diagonal elements $\frac{\partial \varepsilon_{ij}}{\partial t}$ $(i \neq j)$ describing the *shearing* rate of the volume element. The *volume element* is a mathematical abstraction defined as an infinitesimally small cube, but because the smallest unit in brain imaging is a voxel, we may take the voxel as the volume element. Shearing is the deformation that preserves the volume of a voxel but distorts its shape. Note that the sum of the diagonal elements of the strain rate is the first order approximation to the rate of the Jacobian change, i.e.

$$\frac{\partial J}{\partial t} \approx \Lambda_{volume} = \frac{\partial \varepsilon_{11}}{\partial t} + \frac{\partial \varepsilon_{22}}{\partial t} + \frac{\partial \varepsilon_{33}}{\partial t}.$$

It seems that we may have to consider translational, rotational and strain changes for a complete morphological description. However, the most meaningful measurement of brain tissue growth or loss is the rate of the Jacobian change because it directly measures the volumetric changes in the brain. The local translation, the local rotation and the local

shearing change can all be considered as readjustments and reorientations of the local brain-structure due to the volumetric changes in the neighboring regions (Figure 2.3). In between-subject morphological studies of different clinical populations, such measurements might be useful criteria of shape differences. However, in temporally varying within-subject brain morphological studies, we are more interested in regions of brain tissue growth or loss that cause the volumetric changes. Hence, the rate of the Jacobian change is the most meaningful morphological measure of brain tissue growth or loss in deformation-based morphometry.

Finally, the dilatation statistic that consists of spatial derivatives of the displacement field is *statistically independent* from the local translation statistic at each fixed point. To see this, note that any partial derivative of a stationary Gaussian random field is statistically independent from the field itself at a single point [1]. Since the dilatation consists of spatial derivatives of the displacement, it must be statistically independent of the displacement. So the Hotelling's T^2 field of the displacement and the T field of the dilatation measure morphologically different properties at the same voxel.

2.6 Detecting Global Volume Change

Standard MRI-based volumetry, where we are interested in detecting volume changes of the regions of interest (ROI), can be considered as a special case of deformation-based volumetry. Let Ω_t^{ROI} be the 3D region of interest with smooth 2D boundary $\partial\Omega_t^{ROI}$ at time t. The region Ω_0^{ROI} deforms to Ω_t^{ROI} under the deformation $x \to x + U(x,t)$. Note that the volume of Ω_t^{ROI} is given by

$$\|\Omega_t^{ROI}\| = \int_{\Omega_t^{ROI}} dx = \int_{\Omega_0^{ROI}} J(x,t)\, dx.$$

Then the ROI volume-dilatation rate Λ_{ROI} is given by

$$\begin{aligned} \Lambda_{ROI} &= \frac{1}{\|\Omega_0^{ROI}\|}\frac{\partial}{\partial t}\|\Omega_t^{ROI}\| && (2.12)\\ &= \frac{1}{\|\Omega_0^{ROI}\|}\int_{\Omega_0^{ROI}} \frac{\partial J}{\partial t}\, dx && (2.13)\\ &\approx \frac{1}{\|\Omega_0^{ROI}\|}\int_{\Omega_0^{ROI}} \Lambda_{volume}\, dx. && (2.14) \end{aligned}$$

Therefore, the global ROI volume dilatation rate Λ_{ROI} is equivalent to the average of the local volume dilatation rate Λ_{volume} taken over all Ω_0^{ROI}. Since Λ_{volume} is distributed as a

Gaussian random field, Λ_{ROI} becomes a Gaussian random variable. So testing the hypothesis whether there is any volume change between Ω_0^{ROI} and Ω_t^{ROI} can be performed through a simple t-test.

It is also possible to test the global volume change via surface-based deformation analysis. Gauss's *Divergence Theorem* states that

$$\int_{\Omega_0^{ROI}} \nabla \cdot U \; dx = \int_{\partial\Omega_0^{ROI}} U \cdot \mathbf{n} \; dA, \tag{2.15}$$

where $\mathbf{n}$ is a unit normal vector on the surface $\partial\Omega_0^{ROI}$ and dA is the surface area element [75]. It follows that

$$\Lambda_{ROI} \approx \frac{1}{\|\Omega_0^{ROI}\|} \int_{\partial\Omega_0^{ROI}} V \cdot \mathbf{n} \; dA,$$

where $V = \frac{\partial U}{\partial t}$ is the surface displacement velocity on the boundary $\partial\Omega_0^{ROI}$. In Chapter 4, we will develop more sophisticated surface-based local analysis.

2.7 Detecting Global Displacement Change

Instead of testing local translational change, we can test translational change on a global scale. We assume that the displacement vector field is N-dimensional. Suppose that the covariance matrix Σ is known in the linear model (2.2). We are interested in testing the null hypothesis:

$$H_0 : \mu_0(x) = 0 \text{ for all } x \in \Omega.$$

Similar to the Hotelling's T^2 field in (2.3), we define $W(x) = V^t(x)\Sigma^{-1}(x)V(x)$, where V is the displacement velocity. Under H_0, $W(x) = \sum_{i=1}^{N} \epsilon_i^2(x)$ is distributed as a stationary χ_N^2 random field [121], where $\epsilon(x) = (\epsilon_1, \cdots, \epsilon_N)^t$ is the error term defined in (2.2).

Consider another null hypothesis:

$$H_0' : \int_\Omega \|\mu_0\|^2(x) \; dx = 0$$

The two hypothesis H_0 and H_0' are equivalent over the equivalent class of a function g which satisfies $\int_\Omega |g|^2(x) \; dx = 0$. Therefore, instead of testing the null hypothesis H_0, we test H_0'. Under H_0', the exact distribution of the random variable $\int_\Omega W(x) \; dx$ can be found via the *Karhunen-Loève expansion* [2, 38, 66, 126].

Karhunen-Loéve expansion states that for a mean zero Gaussian random field $Z(x)$ with mean square continuity property over a bounded domain $\Omega \subset \mathbb{R}^N$, there exist independent mean zero Gaussian random variables $\{Z_i\}$ and orthonormal bases $\{\phi_i\}$ such that

$$Z(x) = \sum_{i=0}^{\infty} Z_i \phi_i(x). \tag{2.16}$$

Let $\sigma_i^2 = \mathbb{E}Z_i^2 < \infty$ and $R(x, y)$ be the covariance function of $Z(x)$. It can be shown that σ_i^2 and ϕ_i are the i-th eigenvalue and eigenfunction of the integral equation

$$\int_{\Omega} R(x, y)\phi_i(y)\, dy = \sigma_i^2 \phi_i(x). \tag{2.17}$$

Equation (2.17) is a Fredholm equation of the first kind and ϕ_i and σ_i^2 can be estimated numerically if the kernel $R(x, y)$ is given [7]. The bases $\{\phi_i\}$ are orthonormal with respect to the inner product defined by $\langle f, g\rangle = \int_{\Omega} f(x)g(x)\, dx$ such that

$$\langle \phi_i, \phi_j \rangle = \delta_{ij}.$$

The error components ϵ_i are distributed as i.i.d. isotropic Gaussian random fields so they have the orthogonal expansion of the form $\epsilon_i(x) = \sum_{j=0}^{\infty} \epsilon_{ij}\phi_j(x)$ with $\mathbb{E}(\epsilon_{ij}^2) = \sigma_j^2$ and ϵ_{ij} are independent Gaussian random variables for all i, j. Then it follows that

$$\begin{aligned}
\int_{\Omega} W(x)\, dx &= \sum_{i=1}^{N} \langle \epsilon_i, \epsilon_i \rangle \\
&= \sum_{i=1}^{N} \sum_{j,k=0}^{\infty} \int_S \epsilon_{ij}\epsilon_{ik}\phi_j(t)\phi_k(t) dt \\
&= \sum_{i=1}^{N} \sum_{j=0}^{\infty} \epsilon_{ij}^2 \\
&= \sum_{j=0}^{\infty} \sum_{i=1}^{N} \epsilon_{ij}^2.
\end{aligned}$$

Note that

$$\sum_{i=1}^{N} \epsilon_{ij}^2 \overset{D}{\sim} \sigma_j^2 \chi_N^2,$$

where χ_N^2 is the chi-squared distribution with N degrees of freedom. Summing up the above results, we have the exact distribution for global displacement change:

$$\int_S W(x)\, dx \overset{D}{\sim} \sum_{j=0}^{\infty} \sigma_j^2 X_j, \tag{2.18}$$

where $X_j \overset{i.i.d.}{\sim} \chi^2_N$ and $\sigma_j^2 = \int_{\Omega\times\Omega} R(x,y)\phi_j(x)\phi_j(y)\,dxdy$. The sum of the independent chi-squared random variables in (2.18) can be approximated by another chi-square distribution [95]. Note that

$$\begin{aligned}\mathbb{E}\Big(\sum_{j=0}^{\infty}\sigma_j^2 X_j\Big) &= N\sum_{j=0}^{\infty}\sigma_j^2\\ \mathbf{Var}\Big(\sum_{j=0}^{\infty}\sigma_j^2 X_j\Big) &= 2N\sum_{j=0}^{\infty}\sigma_j^4.\end{aligned}$$

Then approximately,

$$\sum_{j=0}^{\infty}\sigma_j^2 X_j \overset{D}{\sim} c\chi^2_\nu,$$

where $c = \sum_{j=0}^{\infty}\sigma_j^4 / \sum_{j=0}^{\infty}\sigma_j^2$ and the approximate degrees of freedom

$$\nu = N\frac{\left(\sum_{j=0}^{\infty}\sigma_j^2\right)^2}{\sum_{j=0}^{\infty}\sigma_j^4}.$$

A similar test procedure that does not use the Karhunen-Loève expansion for testing the null hypothesis of no functional activation can be found in [118].

2.8 Results

Twenty eight normal subjects were selected based on the same physical, neurological and psychological criteria described in [47]. Two T_1-weighted MR scans were acquired for each subject at different times on the same GE Sigma 1.5 T superconducting magnet system. The first scan was obtained at the age 11.5 ± 3.1 years (min. 7.0 yr, max. 17.8 yr) and the second scan was obtained at the age 16.1 ± 3.2 years (min. 10.6 yr, max. 21.8 yr). The time difference between the first and the second scan was 4.6 ± 0.9 years (min. time difference 2.2 yr , max. time difference 6.4 yr). Table 4.1 shows the complete description of the ages. Using the automatic image-processing pipeline [128], a total of 56 MR images were transformed into standardized stereotactic space via a global affine transformation [104] followed by a nonlinear deformation to match the atlas brain Ω_{atlas}. The global affine transformation removes most of the intra- and inter-subject global differences in brain size; adult brains are approximately 5% larger than those of five year old children [35, 36]. Because we are only interested in finding local morphological changes, these global morphological variabilities

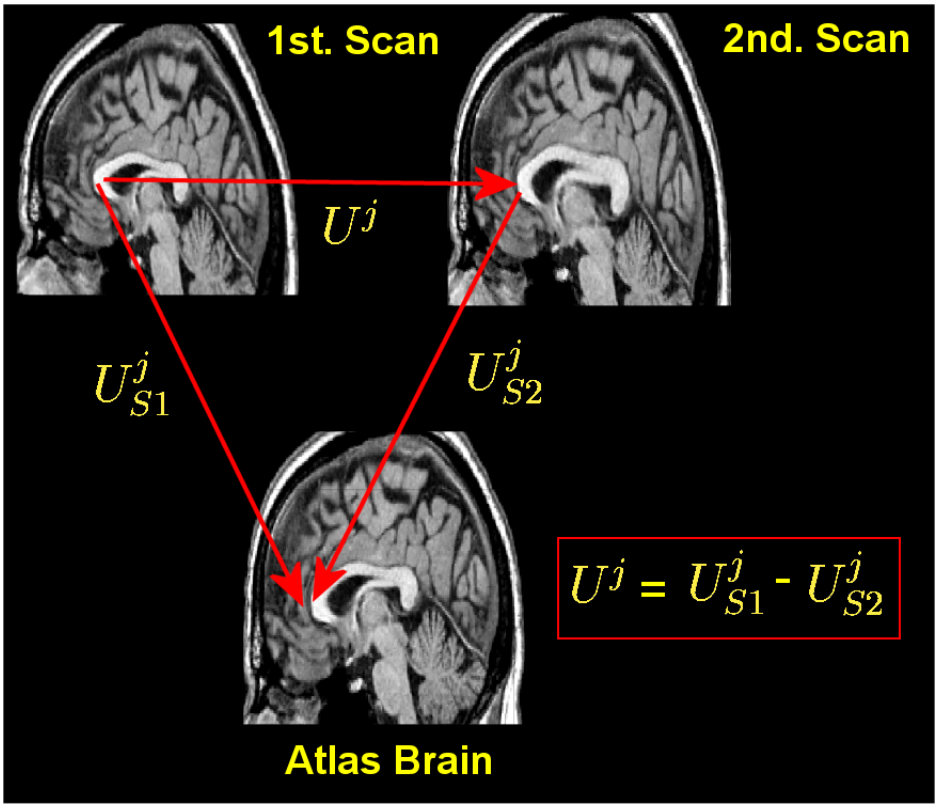

Figure 2.4: The procedure for computing the displacement vector field U^j for subject j. **1.** Compute the displacement field U^j_{S1} from the first scan registered onto the atlas brain Ω_{atlas}. **2.** Compute the displacement field U^j_{S2} from the second scan registered onto the atlas brain Ω_{atlas}. **3.** Compute the difference $U^j = U^j_{S1} - U^j_{S2}$, which is then defined automatically at each voxel $x \in \Omega_{atlas}$.

should be removed via global affine transform in order to improve the power of detection. These registration procedures are based on an automatic multi-resolution intensity matching algorithm [26, 25]. Unlike other registration algorithms that assume a certain fluid dynamics or an elastic deformation model, the intensity-based registration does not assume any explicit physical model in which the deformation from the subject brain to the atlas brain should follow [44, 110]. So the deformation fields obtained from these registration processes can be considered free of any explicit physical model assumption although there might be some intensity-based model assumption,which somehow relates to a physical model.

If U^j_{S1} and U^j_{S2} are the displacement obtained from the non-linear registration of the first and the second scan of subject j to the atlas brain Ω_{atlas} at time t^j_{S1} and t^j_{S2}, the actual displacement U^j between the first scan and the second scan is $U^j = U^j_{S1} - U^j_{S2}$ and the time difference is $t_j = t^j_{S2} - t^j_{S1}$ (Figure 2.4). It is true that if the first scan were directly registered to the second scan without going through the atlas brain, the registration

error would be smaller. However, the displacement fields obtained by the direct registration method still have to be registered onto the atlas brain in order to form statistical parametric maps. The reason for such statistical treatment to analyze the structural data is obvious considering that the displacement field obtained from image registration algorithms for brain development contains a fairly large component of error. The length of the displacement velocity we have observed for the spatially normalized MR scans of 28 normal subjects is usually less 1mm/year, i.e. $\mu_0 = \mathbb{E}\left(\frac{\partial U}{\partial t}\right) \leq 1$ mm/year in average. Optimistically assuming that the image registration algorithm is accurate to within one voxel distance (usually 1 or 2 mm), the registration error seems to be relatively large in brain development. So one may be skeptical about whether the deformation-based morphometry can possibly detect such small changes. Nevertheless it is still possible to pick out the signal when there are enough data; Figure 2.2 illustrates how image smoothing and the statistical treatments improve the power of detection. Statistical treatments compensate for some of such registration errors. Finally the displacement velocity field is smoothed with a 10mm FWHM Gaussian kernel to increase the signal-to-noise ratio (the smoothing parameter FWHM is defined in 3.2). Without the smoothing, it may have been more difficult to detect morphological patterns illustrated in Figure 2.2. However, Gaussian kernel smoothing sometimes tends to blur the fine details of deformation pattern (Figure 2.5).

The regions of statistically significant displacement have been detected (Figure 2.6, yellow) by the Hotelling's T^2 field with the corrected threshold [17]:

$$P\Big(\max_{x\in\Omega_{atlas}} H(x) > 60.0\Big) \approx 0.05.$$

Most of the structural movements have been observed in the frontal lobe *without* any accompanying significant change in local volume. This may indicate that there are continued readjustments of the exact position of brain structures in the frontal lobe without any brain tissue growth or loss in adolescence. Also note that the statistically significant displacement occurs evenly and shows some degree of symmetry between the left and the right hemispheres. Because the local translation statistic measures the relative displacement of brain structure, it does not truly reflect the brain tissue growth process. However, it does indicate the principal direction of the brain growth as shown in the purple box in Figure 2.6 and enlarged in Figure 2.8. Hence, the local translation statistic should be used in conjunction with the local volume change statistic to fully understand the complex dynamics of temporally

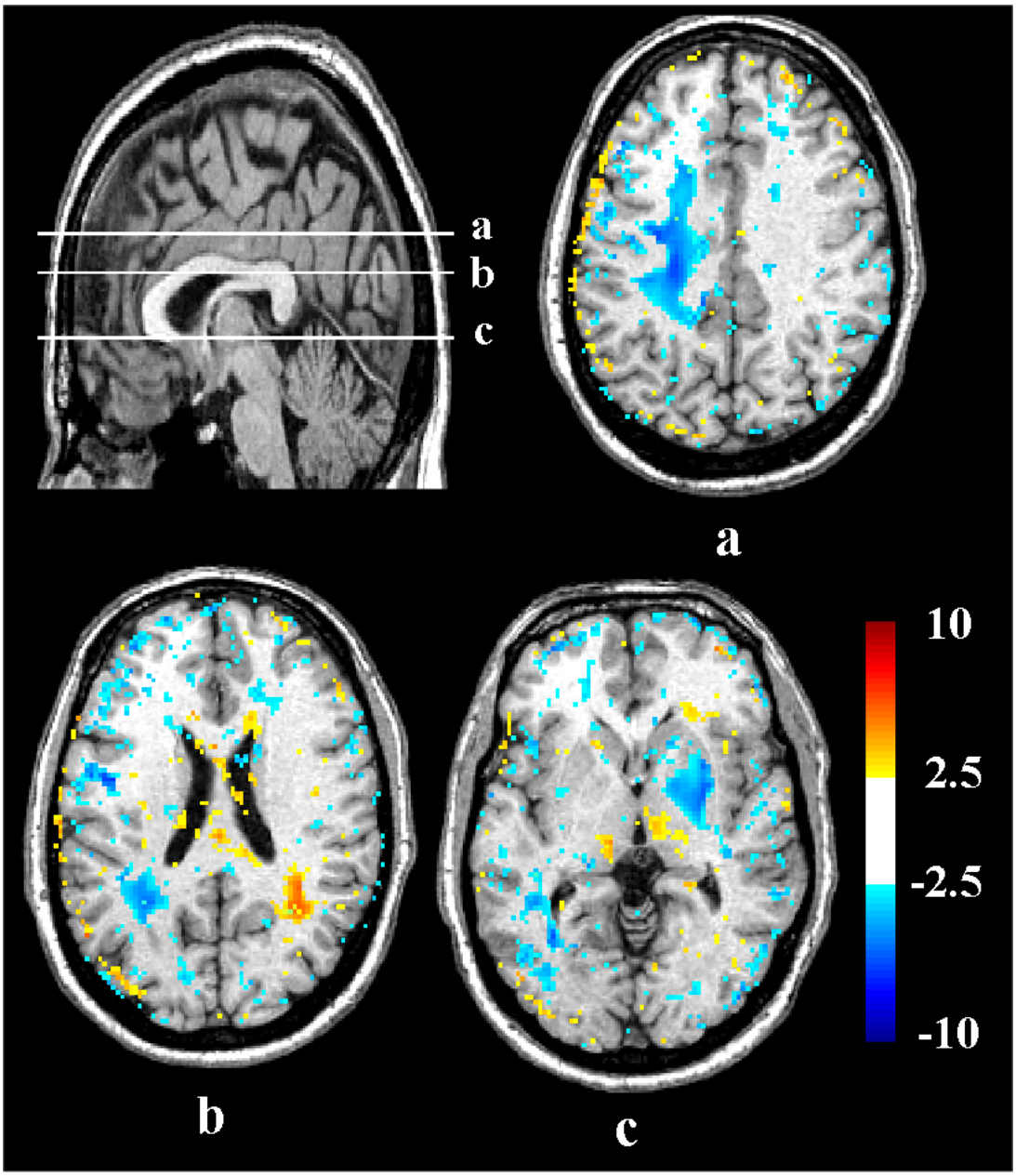

Figure 2.5: t-map of local volume change that is not treated with Gaussian kernel smoothing. Notice that the most of the local volume decrease is concentrated in the white matter. It may be useful as a visual aid for determining deformation patterns, but due to low t-value in most parts, such deformation pattern were found to be statistically insignificant.

changing morphological pattern.

Previous developmental MRI studies have provided evidence for age-related increase in total white matter volume and decrease in total gray matter volume [30, 57, 80, 85, 89] but the analytic procedures used in these studies did not allow the investigators to detect local volume change. The local volume change statistic $T(x)$ is computed using the formula (2.7) with $t_j = t^j_{S1} - t^j_{S2}$. The t-statistic map is thresholded at

$$P\Big(\max_{x\in\Omega_{atlas}} T(x) > 6.5\Big) \approx 0.025,$$

$$P\Big(\max_{x\in\Omega_{atlas}} T(x) < -6.5\Big) \approx 0.025.$$

At this threshold, most of the local volume increase observed around the corpus callosum in Figure 2.2 disappears except for very few localized statistically significant "peaks" in the isthmus and splenium. There was no volume change detected in the rostrum and genu. Figure 2.6 also shows the localized growth in the splenium of the corpus callosum on the coronal section (the single red dot). Therefore, we observe highly focused regions of brain tissue growth at the corpus callosum. [46, 84, 110] reported similar results of growth at the corpus callosum.

The growth at the corpus callosum seems relatively small when compared to the global peaks observed predominantly in somatosensory and motor cortex (the largest red cluster in Figure 2.6). Localized brain tissue loss was also detected at the same time as tissue growth. This tissue loss was highly localized in the subcortical region of the left hemisphere (Figure 2.6, blue). Similar results were also reported in [110], where the extent of the peak growth is wider and less localized than our study has found. It seems our statistical treatments based on the large sample size ($n = 28$) tend to remove a lot of intra-subject variabilities and pick out the common morphological pattern among subjects compared to the smaller sample size ($n = 6$) studied in [110]. Slightly different growth patterns observed between our study and [110] may be due to many factors. Our approach is based on the systematic statistical treatments of large sample size ($n = 28$) with a less accurate intensity based automatic registration algorithm. While the approach taken in [110] is based on a sample size of six without any statistical approach, a more accurate elastic model based registration algorithm with manually matched sulcal landmarks was used. However, the most important difference between the two studies is the age distribution of the subjects. In [110], the age

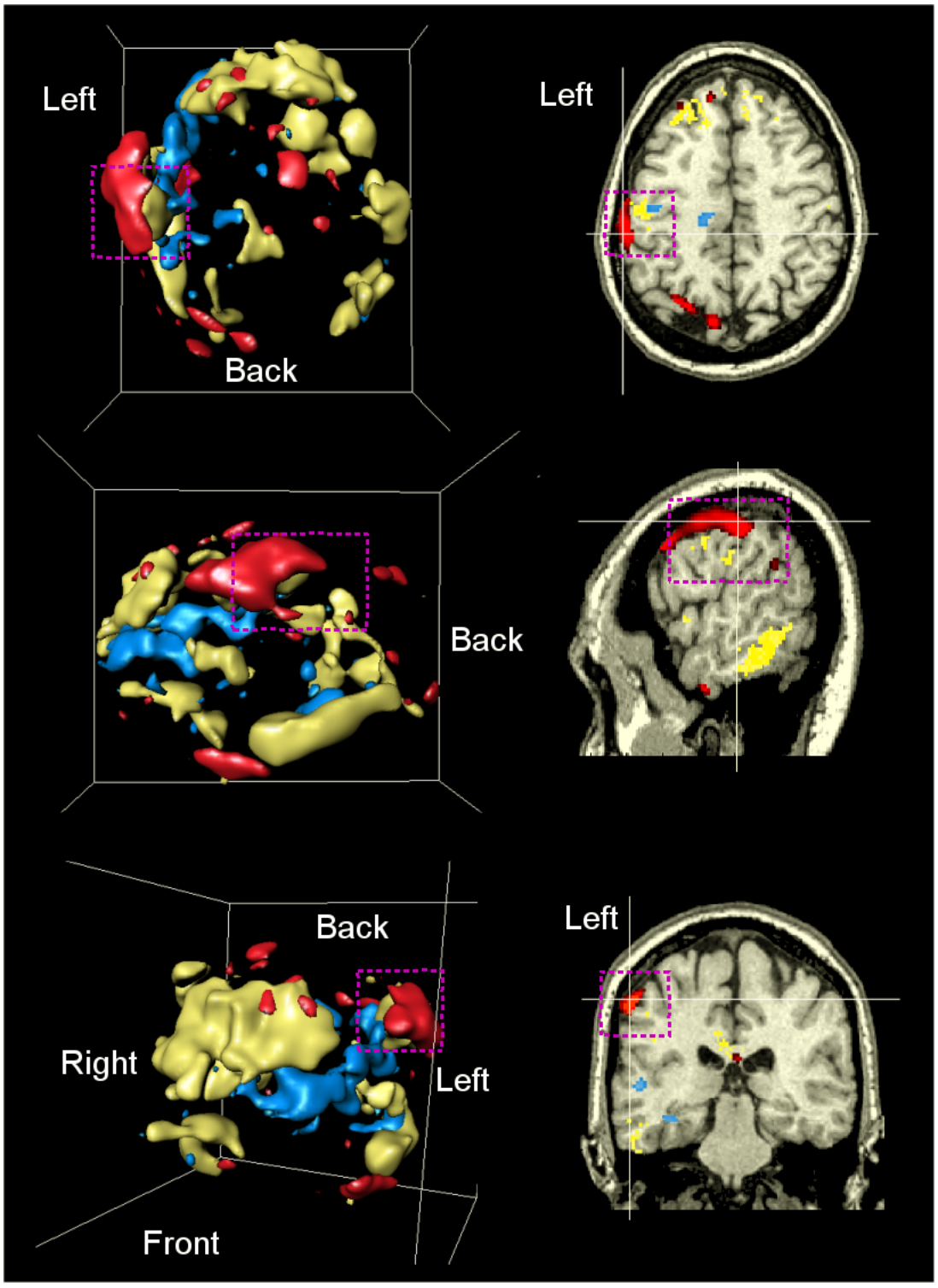

Figure 2.6: **Left**: 3D statistical parametric maps of local volume increase (red), volume decrease (blue) and structural displacement (yellow) thresholded at the probability $0.025, 0.025, 0.05$ (corrected). **Right**: Statistical parametric maps are superimposed on the axial, sagittal and coronal sections of the atlas brain MRI. The cross-sections are taken at the interior of the largest red cluster inside the purple box (parietal cortex). The white lines indicate where the cross-sections are taken.

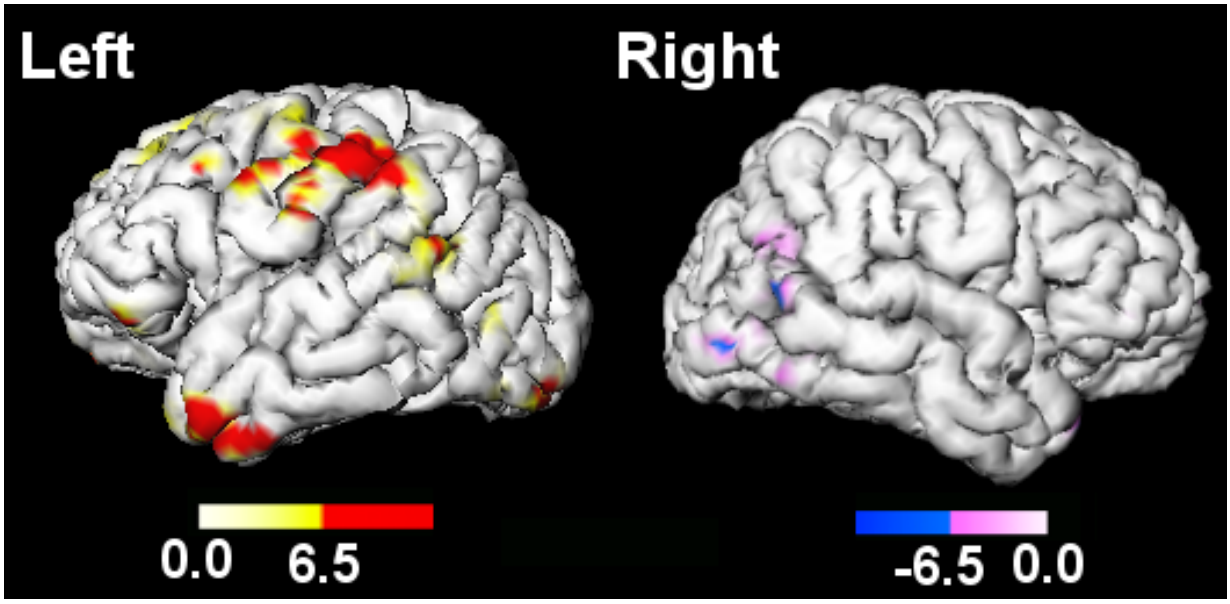

Figure 2.7: 3D statistical parametric map shown in Figure 2.6 is superimposed with the outer cortical surface of the template brain showing dominant local volume increase around the primary motor cortex in the left hemisphere while local volume decrease in the right hemisphere.

distribution of the six subjects is in most part younger than our mean age of 11.5 for the first scan and 16.1 for the second scan. So although there are similar growth patterns common to both studies such as predominant growth at the parietal lobe, localized peak growth at the corpus callosum etc., the two studies are detecting morphological changes in different but nonexclusive age groups.

Finally in answering the question of whether the local translation statistic and the local volume change statistic are measuring different morphological variabilities, we have computed the overlapping regions between the significant volume change and the translation statistics. The volume of the overlapping regions is less than 10% of the total volume of the two statistics combined together. We have already shown that these two statistics are distributed independently at the same voxel. The voxel-by-voxel computation seems to support our claim that these two statistics are indeed measuring different aspects of morphological change. Although they measure different morphological properties, we have observed very interesting relations between these two statistics as illustrated in Figures 2.6 and 2.8. Figure 2.8 is the close-up of the parietal region of the left hemisphere (the purple box in Figure 2.6), showing a large local displacement from the region of volume growth to a region of volume loss, indicating how the structure boundary has moved from the increasing volume to the

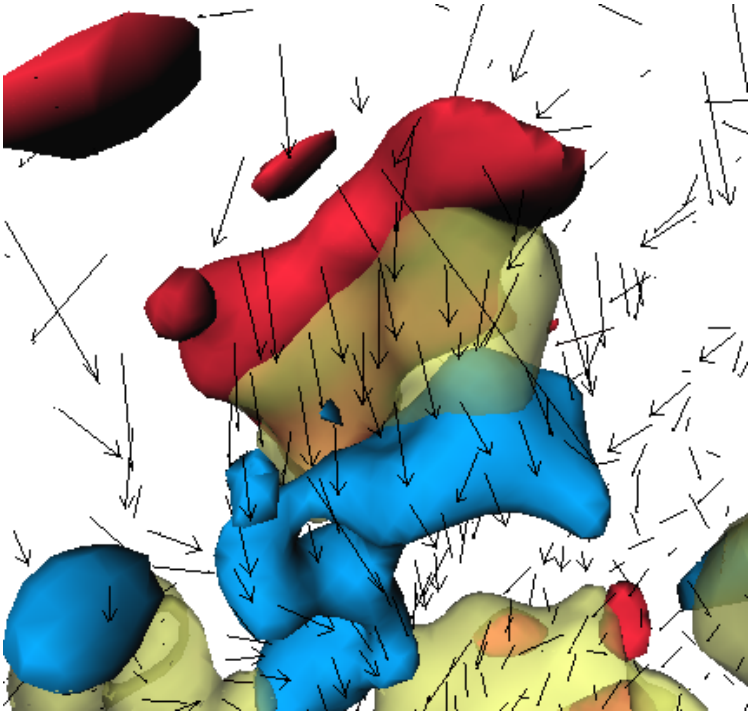

Figure 2.8: A close up of part of the outer left hemisphere inside the purple box in Figure 2.6. Black arrows represent the sample mean displacement velocity subsampled every 10mm and scaled by 50 mm/year. The direction of the mean displacement velocity suggests how the local volume expansion (red) causes the translational movement of the structure (yellow) toward the region of atrophy (blue). The heads of arrows are manually enhanced to clearly indicate the direction of the displacement.

decreasing volume. This phenomenon is also schematically illustrated in Figure 2.3 **b**, where the square grid is undergoing a horizontal translation from the region of volume increase on the left to the region of volume decrease on the right, and Figure 2.3 **d**, where the volume expansion in the middle causes the neighboring structures to radially translate outward. It seems that by studying these two statistical parametric maps simultaneously, the complex dynamic patterns in temporally varying brain morphology can be captured.

Chapter 3

Diffusion Smoothing on Manifolds

We will present two conceptually different approaches to diffusion smoothing on a triangulated cortical surface via the *Laplace-Beltrami operator*, which generalizes an ordinary Laplacian in Euclidean space to manifolds. The first method uses quadratic polynomials for local surface parameterization. Then using a conformal coordinate transform, the Laplace-Beltrami operator is reduced to the planar Laplacian. The second method is based on the *finite element method* of estimating the Laplace-Beltrami operator. As an illustration, the mean curvatures on the outer cortical surface is smoothed to show how the the smoothing incorporates the geodesic curvature information of the surface.

3.1 Diffusion Smoothing

The most general form of Gaussian kernel smoothing of the function $f(x), x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is defined as the convolution of Gaussian kernel K_H with f:

$$F_H(x) = \int_{\mathbb{R}^n} K_H(x - y) f(y)\, dy, \tag{3.1}$$

where $K_H(x) = K(H^{-1}x)/\det(H)$ and the Gaussian kernel is

$$K(x) = \frac{1}{(2\pi)^{n/2}}\, e^{-\|x\|^2/2}.$$

The $n \times n$ symmetric matrix H is called *bandwidth matrix* [96, 100] and it controls the extent of smoothing. Note that $K_H(x)$ is a multivariate normal density with the mean zero and

the covariance matrix H^2. By choosing $H = hI_n$ for some smoothing parameter $h > 0$, we have *isotropic* Gaussian kernel smoothing. The smoothing parameter h controls the extent of smoothing. We will restrict our attention to isotropic Gaussian kernel smoothing and its kernel $K_h(x) \equiv K_{hI_n}(x) = K(x/h)/h^n$. As $h \to 0$, $K_h(x)$ becomes a Dirac delta function $\delta(x)$ [127], which is defined as

$$\int_{\mathbb{R}^n} \delta(x)\, dx = 1,$$
$$\delta(x) = 0 \text{ if } x \neq 0.$$

From the property of the Dirac delta function,

$$\lim_{h \to 0} F_h(x) = \int_{\mathbb{R}^n} \delta(x - y) f(y)\, dy = f(x).$$

So when the bandwidth parameter is too small, we have undersmoothing while when it gets larger, we tend to have oversmoothing. Both oversmoothing and undersmoothing have problems: oversmoothing has small variability but large bias but undersmooting has small bias but large variability. In order to measure over all smoothness, *mean integrated square error* (MISE) has been used as a criterion in most of the theoretical works [15, 96, 100]. The optimal bandwidth parameter h is chosen to minimize MISE. Among imaging researchers, FWHM (*full width at half maximum*) is often used as the smoothing parameter in Gaussian kernel smoothing. FWHM is defined as the the full width at the half maximum of Gaussian kernel (Figure 3.1):

$$\text{FWHM} = 4(\ln 2)^{1/2}\sqrt{t} = 2(2 \ln 2)^{1/2} h. \tag{3.2}$$

Let $t = h^2/2$ and $F(x, t) = F_{\sqrt{2t}}(x)$. We may consider t as time. Using the Fourier transform [41], it can be shown that $F(x, t)$ is the integral solution of the n-dimensional isotropic diffusion equation

$$\frac{\partial F}{\partial t} = \Delta F \tag{3.3}$$

with the initial condition $F(x, 0) = f(x)$, where $\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$ is the Laplacian in n-dimensional Euclidean space. Hence Gaussian kernel smoothing of f is equivalent to the diffusion of the initial data f. The duration of the diffusion of the initial function f determines the extent of smoothing via the relationship $t = h^2/2$. We shall call the smoothing method that is based on solving a diffusion equation as *diffusion smoothing*.

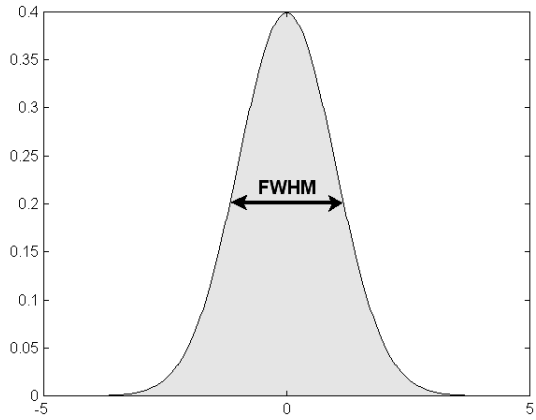

Figure 3.1: FWHM of the Gaussian kernel $K(x)$.

Although diffusion smoothing uses the Laplacian, it is different from *Laplace smoothing*, which penalizes higher powers of the Laplacian Δ in spline smoothing [88]. Because the partial differential equation (PDE) formulation is more adaptable to various situations where the Gaussian kernel smoothing fails, diffusion smoothing has been used in many applications. In anisotropic diffusion, for instance, the extent of smoothing depends on proximity to edges so that it has been mainly used in edge detection [79]. Another application of diffusion smoothing can be found in [86] where it is used to smooth data in a planar region, while constraining the solution to remain within the region via the boundary value problem (BVP) of PDE.

3.1.1 Diffusion Smoothing in $\mathbb{R}$

Let us illustrate diffusion smoothing in $\mathbb{R}$ and shows why it is more robust than Gaussian kernel smoothing. The discrete version of Gaussian kernel smoothing (3.1) on N data points $x^1 < \cdots < x^N \in \mathbb{R}$ is

$$F_h(x) = \frac{x^N - x^1}{N} \sum_{i=1}^{N} K_h(x - x^i) f(x^i), \tag{3.4}$$

which approximates (3.1). Despite the simplicity of Gaussian kernel smoothing, it becomes unstable near the boundary of an area in which the data is defined and this instability worsens as the bandwidth parameter gets larger although it is possible to correct such boundary bias of Gaussian kernel smoothing by using *boundary kernels* [87, 96, 100]. Consider 1-dimensional

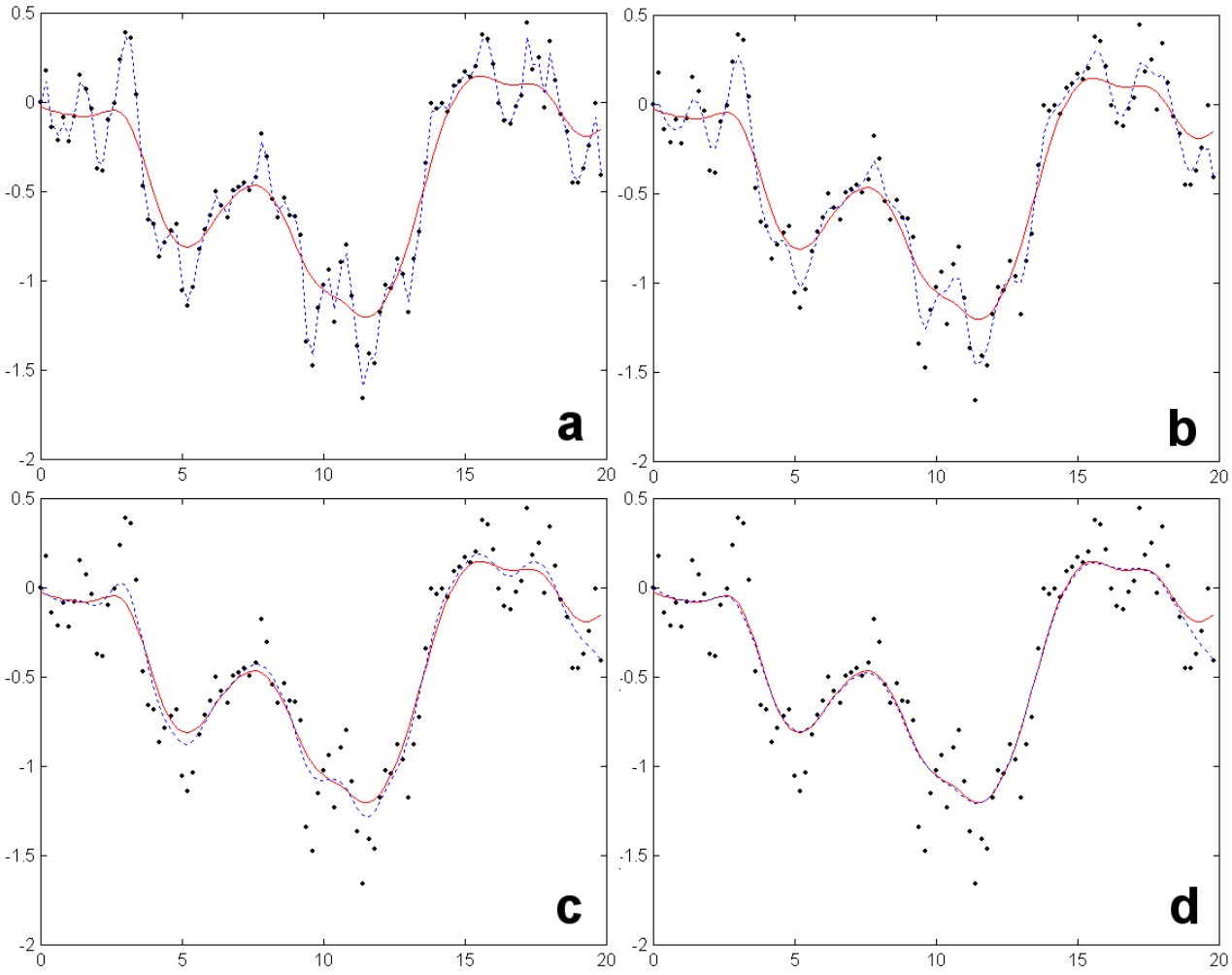

Figure 3.2: Comparison between Gaussian kernel smoothing with the bandwidth parameter $h = 1$ (solid line) and diffusion smoothing (dotted line). **a.** before the iteration. **b.** after 0.05 seconds (5th. iteration). **c.** after 0.25 seconds (25th. iteration). **d.** after 0.5 seconds equivalent to $h = 1$ (50th. iteration).

diffusion equation

$$\frac{\partial F}{\partial t} = \frac{\partial^2 F}{\partial x^2}, \; F(x, 0) = f(x). \tag{3.5}$$

The diffusion equation (3.5) can be solved iteratively by the finite difference method [102]:

$$F(x^i, t_{j+1}) = F(x^i, t_j) + (t_{j+1} - t_j)\widehat{\frac{\partial^2 F}{\partial x^2}}(x^i, t_j), \tag{3.6}$$

where $\widehat{\frac{\partial^2 F}{\partial x^2}}(x^i, t_j)$ is an estimator of $\frac{\partial^2 F}{\partial x^2}(x^i, t_j)$. To simplify the problem, take the same iteration step size $\Delta t = t_{j+1} - t_j$ and iterate until t_j hits $h^2/2$ (Figure 3.2). Estimating the second derivative of a function requires at least three data points. One way of estimating

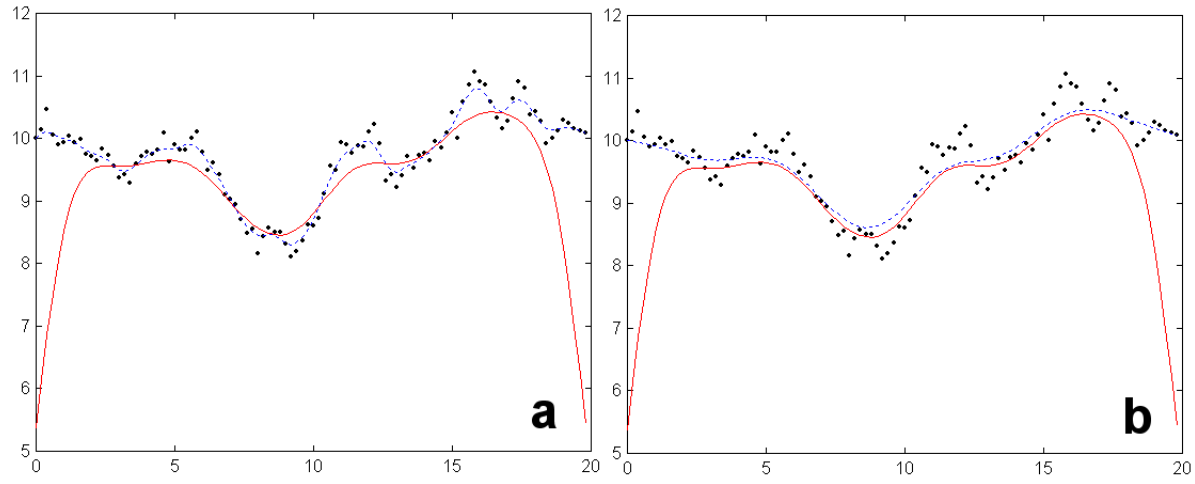

Figure 3.3: Breakdown of Gaussian kernel smoothing with $h = 1$ at the boundary (red line). Diffusion smoothing solved with the boundary condition (blue dotted line). **a.** after 0.25 seconds (25th. iteration). **b.** after 0.5 seconds equivalent to $h = 1$ (50th. iteration).

$\frac{\partial^2 f(x^i)}{\partial x^2}$ is to differentiate a quadratic function that passes through three points

$$\big(x^{i-1}, f(x^{i-1})\big), \big(x^i, f(x^i)\big), \big(x^{i+1}, f(x^{i+1})\big).$$

It can be shown that the estimation based on the quadratic interpolation is

$$\widehat{\frac{\partial^2 f}{\partial x^2}}(x^i) = \frac{\frac{f(x^{i+1})-f(x^i)}{x^{i+1}-x^i} - \frac{f(x^i)-f(x^{i-1})}{x^i-x^{i-1}}}{\frac{x^{i+1}-x^{i-1}}{2}}.$$

We can not interpolate at the end points x^1 and x^N so we are forced to set up boundary conditions $F(x^1, t_j) = f(x^1), F(x^N, t_j) = f(x^N)$ for all t_j in our iteration scheme. With these boundary conditions, we are numerically solving a BVP of PDE and this is why diffusion smoothing will outperform Gaussian kernel smoothing near boundary. Figure 3.3 shows the comparison between Gaussian kernel smoothing in (3.4) and the equivalent diffusion smoothing in (3.6). Note that there are slight discrepancies near the edges due to the fact that the diffusion equation was solved with the boundary conditions while Gaussian kernel smoothing is not.

3.1.2 Smoothing Random Fields

Consider smoothing random noise $X(x)$. We may take $X(x)$ to be a mean zero stationary Gaussian random field in $\mathbb{R}^n$ with the covariance function $R_X(x, y)$. Then Gaussian kernel

smoothing Y of the random field X is defined by

$$Y(x) = \int_{\mathbb{R}^n} K_h(x-y)X(y)\,dy$$

is again a mean zero stationary Gaussian random field with the covariance function

$$R_Y(x,y) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K_h(x-x')K_h(y-y')R_X(x',y')\,dx'\,dy' \tag{3.7}$$

Let $z = x \oplus y$ be a $2n$-dimensional vector formed by stacking two n-dimensional vectors $x = (x^1, \cdots, x^n)^t$ and $y = (y^1, \cdots, y^n)^t$ such that $x \oplus y = (x^1, \cdots, x^n, y^1, \cdots, y^n)^t$. The operator $\oplus$ is usually called the *direct sum*. Gaussian kernel is additive in a sense that $K_h(x)K_h(y) = K_h(x \oplus y)$. Then (3.7) can be rewritten as

$$R_Y(z) = \int_{\mathbb{R}^{2n}} K_h(z-z')R_X(z')\,dz',$$

where $z = x \oplus y$ and $z' = x' \oplus y'$. Hence the covariance function R_Y is Gaussian kernel smoothing of R_X in $2n$-dimension. It is interesting to see that the covariance function R_Y is a solution of a $2n$-dimensional diffusion equation

$$\frac{\partial R}{\partial t} = \Delta R,$$

with the initial condition $R(z,0) = R_X(z)$ after time $t = h^2/2$. In order to find optimal bandwidth parameter h, we minimize a certain penalty function $W_h(x)$. Different penalties will give a different choice of bandwidth. Among digital image researchers, the matched filter theorem [91] has been used as a basis for choosing the optimal filter. The theorem states that the optimal choice of smoothing kernel should match the signal to be detected. For instance, in enhancing sulcal patterns of the outer cortical surface (See Section 3.6), we used 5mm FWHM which is the approximate width of the hollows of the cortical surface along the sulci and visual inspection seems to indicate that 5mm FWHM is a good choice for the sulcal pattern enhancement. Instead of trying to determine the bandwidth parameter in advance, the *scale-space search* method can be used [81, 82, 83, 98, 124]. It searches over a range of bandwidth parameter $h \in [h_0, h_1]$ to find 4D local maxima in both location and parameter space, i.e.

$$\max_{x\in\Omega, h\in[h_0,h_1]} \int_{\mathbb{R}^n} K_h(x-y)X(y)\,dy.$$

Consider the following penalty

$$W_h(x) = \mathbb{E}\,|Y(x) - X(x)|^2 + \mathbf{Var}\ Y(x) \tag{3.8}$$

$$= 2\mathbb{E}\big(Y^2(x)\big) - 2\mathbb{E}\big(Y(x)X(x)\big) + \mathbb{E}\big(X^2(x)\big) \tag{3.9}$$

$$= 2R_Y(x,x) - 2R_{XY}(x,x) + R_X(x,x), \tag{3.10}$$

where R_{XY} is the cross-covariance function of X and Y. When $h \to 0$, the first term $\mathbb{E}\,|Y(x)-X(x)|^2$ vanishes while the second term obtains its maximum $\mathbf{Var}\ Y(x) = R_X(0,0)$. As $t \to \infty$, the diffusion process will reach the stability condition $\frac{\partial R}{\partial t} = 0$ and the covariance function will flatten out to satisfy the harmonic condition $\nabla R = 0$. So when $h \to \infty$, $\mathbf{Var}\ Y(x)$ vanishes while the first term obtains its maximum. Because X and Y are stationary, the penalty function is constant with respect to x, i.e. $W_h(x) = W_h(0)$. But if X is a non-stationary random field, then $W_h(x)$ becomes a function of both x and h. In such a case, the optimal bandwidth h will be different for each x and we get *spatially adaptive smoothing*. For a non-stationary random field X, we can fix the bandwidth by minimizing the integrated penalty over a finite domain Ω, i.e.

$$\min_{h>0} \int_\Omega W_h(x)\ dx.$$

For a stationary random field X, the minimum can be obtained by differentiating (3.10) with respect to h:

$$\frac{\partial R_Y}{\partial h}(x,x) = \frac{\partial R_{XY}}{\partial h}(x,x), \tag{3.11}$$

or equivalently

$$2\mathbb{E}\Big(\frac{\partial Y}{\partial h}(x)Y(x)\Big) = \mathbb{E}\Big(\frac{\partial Y}{\partial h}(x)X(x)\Big). \tag{3.12}$$

The integral equation (3.12) is similar to the Wiener-Hopf equation [38], which is used to find an optimal linear filter and can be solved numerically via the Karhunen-Loève expansion explained in (2.16). Solving (3.12) in general is not easy for a non-trivial covariance function but for a simple case the exact optimal bandwidth can be obtained without resorting to numerical methods. Consider a mean zero stationary Gaussian random field X with the covariance function $R_X(x,y) = \rho K_g(x,y)$ for some constant ρ. We will use the following two identities:

$$K_h(x)K_g(x) = \frac{1}{(2\pi)^{n/2}\beta^n} K_{\alpha/\beta}(x),$$

$$K_h(x)K_g(x-y) = K_{\alpha/\beta}\big(x - \frac{h^2}{\beta^2}y\big)K_\beta(y),$$

where $\alpha = hg$ and $\beta = \sqrt{h^2+g^2}$. From (3.7) and using the fact that $R_Y(x,x) = R_Y(0,0)$,

$$\begin{aligned} R_Y(x,x) &= \rho \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K_h(x)K_h(y)K_g(x-y)\,dx\,dy \\ &= \rho \int_{\mathbb{R}^n} K_\beta(y)K_h(y) \int_{\mathbb{R}^n} K_{\alpha/\beta}(x - \frac{h^2}{\beta^2}y)\,dx\,dy \\ &= \frac{\rho}{(2\pi)^{n/2}\gamma^n}\int_{\mathbb{R}^n} K_{\beta h/\gamma}(y)\,dy \\ &= \frac{\rho}{(2\pi)^{n/2}\gamma^n}, \end{aligned}$$

where $\gamma = \sqrt{h^2+\beta^2} = \sqrt{2h^2+g^2}$. Also

$$\begin{aligned} R_{XY}(x,x) &= \int_{\mathbb{R}^n} K_h(x-y)R_X(x,y)\,dy \\ &= \frac{\rho}{(2\pi)^{n/2}\beta^n}\int_{\mathbb{R}^n} K_{\alpha/\beta}(x-y)\,dy \\ &= \frac{\rho}{(2\pi)^{n/2}\beta^n}. \end{aligned}$$

Solving (3.11), we get

$$h = g\Big(\frac{4^{1/(n+2)}-1}{2-4^{1/(n+2)}}\Big)^{1/2}.$$

When $n = 1,2,3$, we get $h = 1.12g, h = 0.84g, h = 0.69g$ respectively. So once we can estimate the FWHM of the covariance function R_X, it can be used to pick up the optimal bandwidth based on the penalty $W_h(x)$.

Diffusion smoothing has also found applications in the problem of smoothing fMRI data to increase the signal-to-noise ratio on the cortical surfaces of the human brain [4]. When using diffusion smoothing on a curved surface, the smoothing somehow has to incorporate the geometrical features of the curved surface and the Laplacian Δ should change accordingly. For example, on a unit sphere, the spherical Laplacian is

$$\Delta F = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\Big(\sin\theta\frac{\partial F}{\partial\theta}\Big) + \frac{1}{\sin^2\theta}\frac{\partial^2 F}{\partial\phi^2} \tag{3.13}$$

when the spherical coordinates $(x_1,x_2,x_3) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\phi)$ are used. If we use a different coordinate system on the sphere, we will have a different form of the Laplacian since it depends on the coordinate system used. The extension of the Euclidean Laplacian to an arbitrary Riemannian manifold is called the *Laplace-Beltrami operator* [51, 64]. In the

following sections, we will explicitly explain two computational methods for estimating the Laplace-Beltrami operator on triangulated meshes. The first, which we call the parametric method uses a local quadratic parameterization of the surface, and then a transformation to an orthogonal coordinate system on which the Laplace-Beltrami operator will be reduced to the simple 2-dimensional Euclidean Laplacian. The other approach uses a finite element formulation and provides a discrete Laplace-Beltrami operator using only the edge lengths of the triangulated surface.

3.2 Differential Geometry

Before we introduce diffusion smoothing on an arbitrary manifold, let us review some basic differential geometry relevant to the Laplace-Beltrami operator and curvature. See classical books by W.M. Boothby [14], M.P.D. Carmo [18] and E. Kreyszig [64] for an overview of differential geometry.

Suppose we have an orientable surface $\partial\Omega$ which we assume to be a smooth twice-differentiable 2-dimensional manifold embedded in $\mathbb{R}^3$. Then we have a parameterization of the surface $\partial\Omega$:

$$X(u) = \{x_1(u), x_2(u), x_3(u) : u = (u^1, u^2) \in D\}$$

where all partial derivatives of X up to the second order are continuous in a planar domain D. The smooth map $X : D \to \partial\Omega$ is called a *parameterized surface* of $\partial\Omega$ if the partial derivative vectors

$$X_1(u) = \left(\frac{\partial x_1}{\partial u^1}, \frac{\partial x_2}{\partial u^1}, \frac{\partial x_3}{\partial u^1}\right)^t \quad \text{and} \quad X_2(u) = \left(\frac{\partial x_1}{\partial u^2}, \frac{\partial x_2}{\partial u^2}, \frac{\partial x_3}{\partial u^2}\right)^t \tag{3.14}$$

form a basis for the tangent plane $T_{\mathbf{p}}(\partial\Omega)$ at any $p = X(u) \in \partial\Omega$, i.e. $X_1(u) \times X_2(u) \neq 0$ for any $u \in D$.

3.2.1 Riemannian Metric Tensors

Because X_1 and X_2 form a basis for the tangent plane $T_{\mathbf{p}}(\partial\Omega)$, any vector $d\xi \in T_{\mathbf{p}}(\partial\Omega)$ can be written as a linear combination of the basis vectors X_1 and X_2: $d\xi = du^1 X_1 + du^2 X_2$ for

some constants du^1 and du^2. Then the length of the vector $d\xi$ in the Cartesian coordinate is

$$d\xi^2 \equiv \langle d\xi, d\xi \rangle = g_{11} du^1 du^1 + g_{12} du^1 du^2 + g_{21} du^2 du^1 + g_{22} du^2 du^2, \tag{3.15}$$

where the coefficients $g_{ij} = \langle X_i, X_j \rangle$ are called the *Riemannian metric tensor* and they measure the amount of deviation from the Cartesian coordinate system. The bilinear form (3.15) is called the *first fundamental form*. The first fundamental form enables us to compute intrinsic properties of the surface such as lengths, angles and areas. If a curve on the surface $\partial\Omega$ is given by $X(u(s))$, where the curvilinear coordinates $u(s) = (u^1(s), u^2(s))$ is parameterized by a single parameter s, its length is given by

$$\int \|\dot{u}\| \, ds = \int \Big(\sum_{i,j} g_{ij} \frac{\partial u^i}{\partial s} \frac{\partial u^j}{\partial s} \Big)^{1/2} ds.$$

The angle θ between two vectors $\xi, \eta \in T_{\mathbf{p}}(\partial\Omega)$ can be computed in terms of the Riemannian metric tensor in the following way:

$$\cos\theta = \frac{\xi \cdot \eta}{\|\xi\| \|\eta\|} = \frac{\sum_{i,j} g_{ij} \xi^i \eta^j}{(\sum_{i,j} g_{ij} \xi^i \xi^j)^{1/2} (\sum_{i,j} g_{ij} \eta^i \eta^j)^{1/2}}.$$

The total surface area of a region $A \subset \partial\Omega$ is

$$\int_{X^{-1}(A)} \sqrt{\det g} \, du^1 du^2,$$

where $\det g = g_{11} g_{22} - g_{12}^2$ and does not depends on the parameterization X [18, pp. 97]. Note that $\sqrt{\det g}$ is the local surface area element, which will be used in Chapter 4 in measuring the amount of local surface area change.

3.2.2 Gaussian and Mean Curvatures

The unit outward normal vector $\mathbf{n}$ to the surface is given by

$$\mathbf{n} = \frac{X_1 \times X_2}{\sqrt{\det g}}.$$

Then the vectors $(X_1, X_2, \mathbf{n})$ form an orthogonal basis in the 3D Euclidean space. So the partial derivative of the basis vectors X_1, X_2 and $\mathbf{n}$ can be expressed in terms of the basis:

$$\begin{aligned} \frac{\partial X_i}{\partial u^j} &= \Gamma_{ij}^1 X_1 + \Gamma_{ij}^2 X_2 + \Gamma_{ij}^3 \mathbf{n}, \\ \frac{\partial \mathbf{n}}{\partial u^j} &= \Gamma_{3j}^1 X_1 + \Gamma_{3j}^2 X_2 + \Gamma_{3j}^3 \mathbf{n}, \end{aligned}$$

where the constants Γ_{ij}^k are called the Christoffel symbols.

The *second fundamental form* is given by

$$-\langle d\xi, d\mathbf{n}\rangle = l_{11}du^1du^1 + l_{12}du^1du^2 + l_{21}du^2du^1 + l_{22}du^2du^2, \tag{3.16}$$

where $l_{ij} = \langle X_{ij}, \mathbf{n}\rangle$ and

$$X_{ij} = \frac{\partial X_i}{\partial u^j} = \Big(\frac{\partial^2 x_1}{\partial u^i \partial u^j}, \frac{\partial^2 x_2}{\partial u^i \partial u^j}, \frac{\partial^2 x_3}{\partial u^i \partial u^j}\Big)^t.$$

Let $g = (g_{ij})$ and $l = (l_{ij})$. The *principal curvatures* κ_1 and κ_2 are defined as the eigenvalues of $g^{-1}l$ and the mean curvature K_M and the Gaussian curvature K_G can be given in terms of the principal curvatures as

$$K_M = (\kappa_1 + \kappa_2)/2 = \mathrm{tr}(g^{-1}l)/2 \text{ and } K_G = \kappa_1\kappa_2 = \det(l)/\det(g). \tag{3.17}$$

In estimating the Gaussian curvature on a triangulated mesh, there is an alternative method based on the covariance matrix of the surface normals [5]. There is also a finite element version of the mean curvature estimation [37, 78].

For the surface of the form

$$x_3 = z(x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \frac{1}{2}\beta_3 x_1^2 + \beta_4 x_1 x_2 + \frac{1}{2}\beta_5 x_2^2 + \cdots,$$

we can parameterize it by $X(u^1, u^2) = \big(u^1, u^2, z(u^1, u^2)\big)$. A simple computation shows that $X_1 = (1, 0, \beta_1), X_2 = (0, 1, \beta_2), X_{11} = (0, 0, \beta_3), X_{12} = (0, 0, \beta_4), X_{22} = (0, 0, \beta_5)$ at the origin. Then the normal vector $\mathbf{n}$ is

$$\mathbf{n} = \frac{(-\beta_1, -\beta_2, 1)}{(1 + \beta_1^2 + \beta_2^2)^{1/2}}.$$

From (3.15) and (3.16), the coefficients of the fundamental forms at the origin are given by

$$g = \begin{pmatrix} 1 + \beta_1^2 & \beta_1\beta_2 \\ \beta_1\beta_2 & 1 + \beta_2^2 \end{pmatrix}, \qquad l = \frac{1}{(1 + \beta_1^2 + \beta_2^2)^{1/2}} \begin{pmatrix} \beta_3 & \beta_4 \\ \beta_4 & \beta_5 \end{pmatrix}. \tag{3.18}$$

Then the mean curvature is given by

$$K_M = \frac{\beta_3(1 + \beta_2^2) + \beta_5(1 + \beta_1^2) - 2\beta_1\beta_2\beta_4}{(1 + \beta_1^2 + \beta_2^2)^{3/2}}. \tag{3.19}$$

So it is easy to compute geometric quantities such as local surface area, length and curvatures when the local quadratic surface patch is used. Because of this simplicity, quadratic surface fitting has been used in estimating curvatures of a macaque brain surface [60].

3.2.3 Laplace-Beltrami Operator

The gradient ∇_X of F on the tangent plane $T_{\mathbf{p}}(\partial\Omega)$ is defined as

$$\nabla_X F = \sum_{i,j} g^{ij} \frac{\partial F}{\partial u^j} X_i, \tag{3.20}$$

where $(g^{ij}) = g^{-1}$ [75, p. 69]. Then the generalized Laplacian called the *Laplace-Beltrami operator* Δ_X corresponding to the surface parameterization X is defined as the divergence of the gradient operator such that

$$\Delta_X F = \nabla_X \cdot (\nabla_X F) = \frac{1}{|g|^{1/2}} \sum_{i,j} \frac{\partial}{\partial u^i} \Big(|g|^{1/2} g^{ij} \frac{\partial F}{\partial u^j} \Big) \tag{3.21}$$

[64, 75]. For the derivation of the Laplace-Beltrami operator without using differential geometry, one may approach the problem in terms of a curvilinear coordinate transform [29]. One of the most important properties of the Laplace-Beltrami operator is that it is independent of the parameterization of $\partial\Omega$, i.e. if $\widetilde{X} = X \circ \Phi$ is another parameterization of the surface $\partial\Omega$,

$$\Delta_{\widetilde{X}} \widetilde{F} = \widetilde{\Delta_X F}.$$

However, this equivalence may fail numerically when we estimate ΔF with different parameterizations, so that great care should be taken to choose a proper parameterization which stabilizes the numerical computation and minimizes the variances of errors in estimating the Laplace-Beltrami operator. Another important property of the Laplace-Beltrami operator, which will be used in the finite element method, is that the operator is *self-adjoint*. If F and G are twice differentiable functions on $\partial\Omega$, then

$$\int_{\partial\Omega} G \Delta F \, dS = \int_{\partial\Omega} F \Delta G \, dS = - \int_{\partial\Omega} \langle \nabla F, \nabla G \rangle \, dS,$$

where the inner product

$$\langle \nabla F, \nabla G \rangle = \sum_{ij} g^{ij} \frac{\partial F}{\partial u^i} \frac{\partial F}{\partial v^i}$$

and the surface area element $dS = \sqrt{\det g} \, du^1 du^2$ [51, pp.143].

Conformal coordinates are defined as a coordinate system $u = (u^1, u^2)$ whose metric is given by $d\xi^2 = \lambda(du^1)^2 + \lambda(du^2)^2$ for some function $\lambda = \lambda(u)$. With respect to the conformal coordinates, the Laplace-Beltrami operator can be simplified to

$$\Delta_X = \frac{1}{\lambda} \Big(\frac{\partial^2}{\partial (u^1)^2} + \frac{\partial^2}{\partial (u^2)^2} \Big).$$

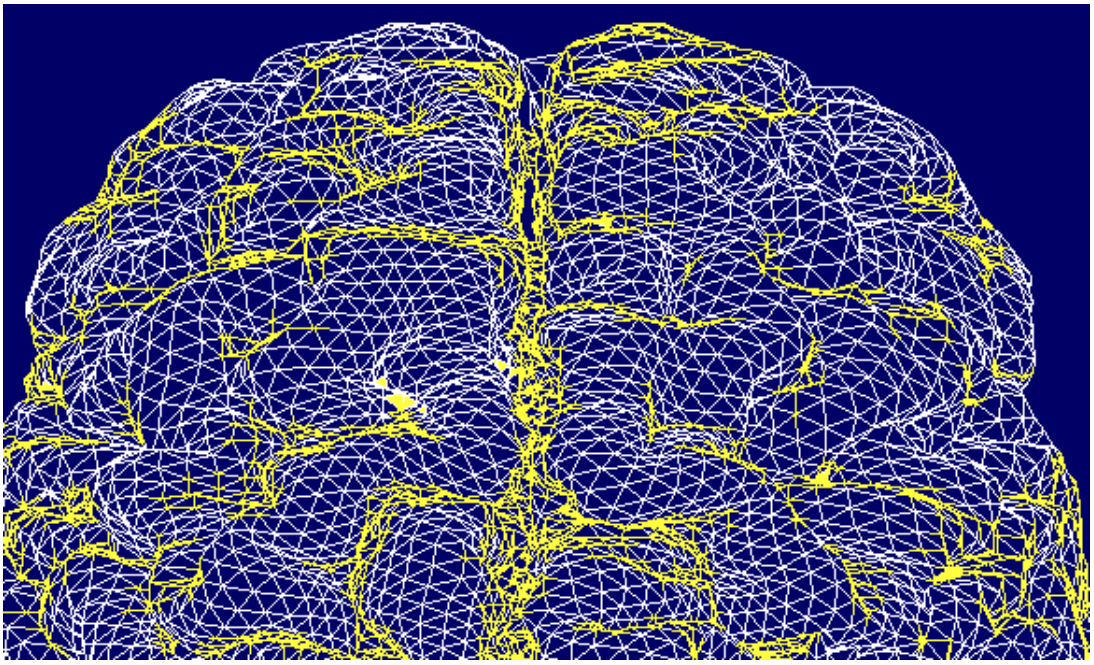

Figure 3.4: A typical triangular mesh of the outer cortical surface consisting of 81,920 triangles and 40,962 vertices.

For an arbitrary smooth surface and a fixed point $\mathbf{p}$, we can always find conformal coordinates such that $X(u) = \mathbf{p}$ and $\lambda(u) = 1$ [18]. Therefore, if we find a conformal coordinate system at each $\mathbf{p} \in \partial\Omega$, the computation of the Laplace-Beltrami operator at $\mathbf{p} = X(u)$ becomes the planar Laplacian at u. In the following section, we will explain a way to find such conformal coordinate systems on a triangulated surface by a simple affine transformation of the coordinates.

3.3 Parametric Method

The standard method for triangulating the surface is the marching cubes algorithm [72]. Alternative methods such as the level set method [97] or deformable surfaces method [33] are available. We have used the anatomic segmentation using proximities (ASP) method [74], which is a variant of the deformable surfaces method. In triangulating cortical surfaces of the human brain from 3D MRI, we have used the ASP method to extract 81,920 triangles. At this surface sampling rate, the average intervertex distance is about 3mm (Figure 3.4).

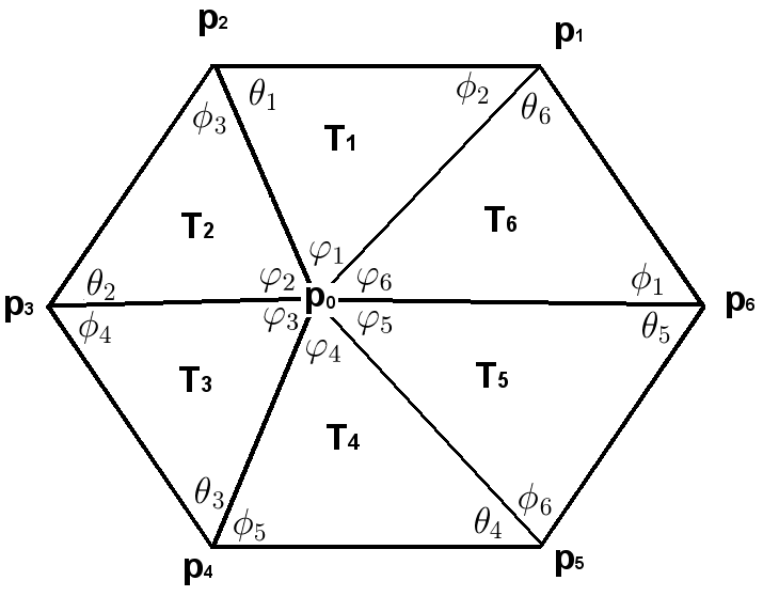

Figure 3.5: A typical triangulation in the neighborhood of $\mathbf{p} = \mathbf{p}_0$. The triangulation is either pentagonal or hexagonal in ASP algorithm for 81920 triangles.

3.3.1 Estimating Normal Vectors

In order to compute the Riemannian metric tensors on the triangulated surface, we first estimate the tangent plane and its normal vector at each node then find a local parameterization in the neighborhood of each node. Usually normal vectors are computed during the triangulation process. In the ASP method, the outward unit normal vector $\mathbf{n}$ at each node p is computed as the weighted average of the unit normals of the incident triangles. If $\mathbf{p}_1, \ldots, \mathbf{p}_m$ are m neighboring points of $\mathbf{p} = \mathbf{p}_0$ in the counter-clockwise direction with respect to the tangent plane $T_{\mathbf{p}}(\partial\Omega)$ at $\mathbf{p}$ (Figure 3.5), then the unit normal vector $\mathbf{n}$ is estimated as

$$\mathbf{n} = \frac{\sum_{i=1}^{m} \varphi_i \mathbf{n}_i}{\sum_{i=1}^{m} \varphi_i},$$

where the unit vectors $\mathbf{n}_i$ are normal to each triangle T_i.

$$\mathbf{n}_i = \frac{(\mathbf{p}_{i+1} - \mathbf{p}) \times (\mathbf{p}_i - \mathbf{p})}{\|(\mathbf{p}_{i+1} - \mathbf{p}) \times (\mathbf{p}_i - \mathbf{p})\|}$$

and the interior angles are

$$\varphi_i = \cos^{-1} \frac{\langle \mathbf{p}_{i+1} - \mathbf{p}, \mathbf{p}_i - \mathbf{p} \rangle}{\|\mathbf{p}_{i+1} - \mathbf{p}\| \|\mathbf{p}_i - \mathbf{p}\|}.$$

Alternatively, we may employ a method similar to principal components analysis (PCA). The equation of the plane with the unit normal vector $\mathbf{n}$ passing through the point p is $\langle \mathbf{n}, x \rangle = \langle \mathbf{n}, \mathbf{p} \rangle$. The distance from the point $\mathbf{p}_i$ to the plane is the length of the projection of $\mathbf{p}_i - \mathbf{p}$ onto the unit normal vector $\mathbf{n}$, i.e. $\langle \mathbf{n}, \mathbf{p}_i - \mathbf{p} \rangle$. Then we find the best fitting tangent plane in the sense of minimizing the sum of squared distance of the points $\mathbf{p}_1, \ldots, \mathbf{p}_m$ to the plane:

$$\min_{\mathbf{n}} \sum_{i=1}^{m} \langle \mathbf{n}, \mathbf{p}_i - \mathbf{p} \rangle^2 = \min_{\mathbf{n}} \ \mathbf{n}^t \mathbf{C} \mathbf{n},$$

where $\mathbf{C} = \sum_{i=1}^{m} (\mathbf{p}_i - \mathbf{p})(\mathbf{p}_i - \mathbf{p})^t$. If the fitting plane is not forced to pass through the point p, $\mathbf{C}$ becomes the sample covariance matrix of $\mathbf{p}_1, \ldots, \mathbf{p}_m$ and the optimization problem is exactly the standard PCA. Since $\mathbf{n}^t \mathbf{n} = 1$, using the Lagrange multiplier γ minimize $\mathbf{n}' \mathbf{C} \mathbf{n} - \gamma(\mathbf{n}' \mathbf{n} - 1)$. Differentiating with respect to $\mathbf{n}$, $\mathbf{C}\mathbf{n} - \gamma \mathbf{n} = 0$. Thus, γ is an eigenvalue of $\mathbf{C}$. Note that we are minimizing $\mathbf{n}' \mathbf{C} \mathbf{n} = \mathbf{n}' \gamma \mathbf{n} = \gamma$. So the unit normal vector $\mathbf{n}$ of the best fitting tangent plane should be the eigenvector $\mathbf{n}$ that corresponds to the smallest eigenvalue.

If the interior angle of triangles joining the vertex $\mathbf{p}$ is acute, the best fitting plane passing through $\mathbf{p}$ might end up perpendicular to the tangent plane $T_{\mathbf{p}}(\partial\Omega)$. In this case, the unit normal vector $\mathbf{n}$ to $T_{\mathbf{p}}(\partial\Omega)$ should be the eigenvector that corresponds to the largest eigenvalue.

3.3.2 Global Parameterization: B-Splines

The ASP algorithm provides a one-to-one mapping from the surface to a sphere [74]. Using *spherical projection*, every point except the north pole on the sphere can be mapped onto the plane. Spherical projection has been used as a global parameterization in flattening the whole brain surface onto a single plane [6]. One of the most widely used surface parameterization method is to extend the B-Spline curves to surfaces via tensor product. Let us review special B-spline curves called *Bézier curves* [52, 76]. Consider $m+1$ given points $\mathbf{p}_0, \cdots, \mathbf{p}_m \in \mathbb{R}^N$. The simplest Bézier curve is a line segment $\mathbf{p}_{i,j}(u)$ joining two points $\mathbf{p}_i$ and $\mathbf{p}_j$:

$$\mathbf{p}_{i,j}(u) = (1-u)\mathbf{p}_i + u\mathbf{p}_j, \ u \in (0,1).$$

For three points $\mathbf{p}_0, \mathbf{p}_1$ and $\mathbf{p}_2$, a quadratic Bézier curve is given by

$$\mathbf{p}_{0,2}(u) = (1-u)\mathbf{p}_{0,1} + u\mathbf{p}_{1,2} = (1-u)^2\mathbf{p}_0 + 2(1-u)\mathbf{p}_1 + u^2\mathbf{p}_2.$$

In general, m-th order Bézier curve is

$$\mathbf{p}_{0,m}(u) = \sum_{j=0}^{m} \phi_{j,m}(u)\mathbf{p}_j, \ u \in (0,1)$$

where the basis functions $\{\phi_{j,m}(u) = \binom{m}{j}(1-u)^{m-j}u^j, j = 0, \ldots, m\}$ are Bernstein polynomials of degree m. Cubic Bézier curves are the most often used Bézier curve because cubics satisfy the minimum curvature or strain energy property, which make them a more suitable tool for a smooth curve approximation [52, 76]. To avoid increasing the degree of the Bézier curve, we need to piece together Bézier curves. If continuity conditions are satisfied for each Bézier curve segments, the result is a *B-spline* curve. In general, a B-spline curve of degree $K-1$ with $m+1$ vertices $\mathbf{p}_0, \cdots \mathbf{p}_m$ is defined as

$$X(u) = \sum_{j=0}^{m} B_{j,K}(u)\mathbf{p}_j,$$

where the B-spline functions $\{B_{j,K}(u), j = 0, \cdots, m\}$ are defined recursively in [52, 76]. Then we can use these B-spline curves to generate surfaces. Consider a rectangular mesh of vertices $\{V_{ij}\}$. A B-spline surface parameterization X of degree $K-1$ for points $\mathbf{p}_{ij} \in \mathbb{R}^3$ can be defined by the tensor product:

$$X(u^1, u^2) = \sum_{i,j} B_{i,K}(u^1)B_{j,K}(u^2)\mathbf{p}_{ij}. \tag{3.22}$$

The advantage for using B-spline to represent the surface is that it is easy to evaluate the curvature of a surface or other geometric characteristics of the surface because polynomial functions can be differentiated easily. The disadvantage of using the tensor B-spline is that it is not easy to modify the above formulation which works so well for a rectangular mesh to a irregular triangular mesh.

3.3.3 Local Parameterization: Quadratic Surface

Instead of using B-splines to form a parametric surface, there is a simpler method based on the *polynomial regression* [90]. This is a smoothing technique to fit the given points

$\mathbf{p}_0, \ldots, \mathbf{p}_m$ by the least-squares method to a polynomial function of the form

$$f(x,y) = \sum_{i+j \le p} \beta_{ij}\, x^i y^j.$$

Then for $\mathbf{p}_i = (x^i, y^i, z^i)$, $P = \frac{(p+1)(p+2)}{2}$ unknown coefficients β_{ij}'s are chosen to minimize the residual

$$\sum_{i=0}^{m} \left[z^i - f(x^i, y^i)\right]^2.$$

The drawback of the polynomial regression is that there is a tendency to weave the outer most vertices to find vertices in the center. Therefore, polynomial regression is not recommended for global surface parameterization. Our surface-based morphometry will try to avoid using any global surface parameterization. Other families of surface parameterizations have been suggested but are rarely used. One of them is to use finite Fourier series:

$$\begin{aligned} f(x,y) \;=\; & \sum_{0 \le i,j \le r} [a_{ij} \sin(i\omega_1 x)\sin(j\omega_2 y) + b_{ij}\sin(i\omega_1 x)\cos(j\omega_2 y) \\ & + c_{ij}\cos(i\omega_1 x)\sin(j\omega_2 y) + d_{ij}\cos(i\omega_1 x)\cos(j\omega_2 y)]. \end{aligned}$$

Such a surface requires a large number of coefficients $a_{ij}, b_{ij}, c_{ij}, d_{ij}$ plus the fundamental frequencies ω_1, ω_2 to be estimated. Hence, surface fitting based on finite Fourier series on the large data set is computationally intensive although the fit would be better than the standard polynomial regression.

However, in estimating the Laplace-Beltrami operator or curvatures, it is not necessary to find such global parameterization of the surface $\partial\Omega$. A local surface parameterization in the neighborhood of $\mathbf{p}$ can be obtained via the projection of the local surface onto the tangent plane $T_{\mathbf{p}}(\partial\Omega)$. Let Q be an orthogonal matrix which rotates the normal vector $\mathbf{n}$ to align with the x_3 axis, i.e. $Q\mathbf{n} = (0,0,1)^t$. It is easy to see that

$$Q = \begin{pmatrix} n_3 & 0 & -\sqrt{n_1^2+n_2^2} \\ 0 & 1 & 0 \\ \sqrt{n_1^2+n_2^2} & 0 & n_3 \end{pmatrix} \begin{pmatrix} \frac{n_1}{\sqrt{n_1^2+n_2^2}} & \frac{n_2}{\sqrt{n_1^2+n_2^2}} & 0 \\ -\frac{n_2}{\sqrt{n_1^2+n_2^2}} & \frac{n_1}{\sqrt{n_1^2+n_2^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is such an orthogonal matrix assuming $n_1, n_2 \neq 0$. If $n_1 = n_2 = 0$, we can take $Q = I_3$. Let $x \in \partial\Omega$ be a point in the neighborhood of $\mathbf{p}$. Consider the transformation defined by $y = (y_1, y_2, y_3)^t = Q(x - \mathbf{p})$. Under this transformation, the local surface region translates to

the origin and then rotates by Q. Then with respect to the new coordinates (y_1, y_2, y_3), the local surface can be explicitly written as $y_3 = z(y_1, y_2)$ for some function z assuming local smoothness of the surface. Hence by identifying (y_1, y_2) as our parameter space, we have the following local parameterization in the neighborhood of $\mathbf{p}$:

$$X(u^1, u^2) = \mathbf{p} + Q^t\Big(u^1, u^2, z(u^1, u^2)\Big)^t. \tag{3.23}$$

Then the basis on the tangent plane is

$$X_1 = Q^t\Big(1, 0, \frac{\partial z}{\partial u^1}|_{(0,0)}\Big)^t \text{ and } X_2 = Q^t\Big(0, 1, \frac{\partial z}{\partial u^2}|_{(0,0)}\Big)^t.$$

Thus the Riemannian metric tensor is given by

$$g_{ij} = \langle X_i, X_j\rangle = \delta_{ij} + \frac{\partial z}{\partial u^i}\frac{\partial z}{\partial u^j}, \tag{3.24}$$

where the derivatives are evaluated at $(0, 0)$. Hence the metric tensor at $\mathbf{p}$ is completely determined by the derivatives of the function z evaluated at $(0, 0)$ and it is independent of the rotation of the tangent plane by Q. Similarly, the coefficients of the second fundamental form are invariant under such a transformation.

In the neighborhood of $(0, 0)$, we have the Taylor approximation of the function z:

$$z(u^1, u^2) = \beta_1 u^1 + \beta_2 u^2 + \beta_3 (u^1)^2 + \beta_4 u^1 u^2 + \beta_5 (u^2)^2 + \cdots. \tag{3.25}$$

Since we are forcing the function z to pass through the origin, there is no constant term in the Taylor expansion. The problem of estimating the coefficients β_i can be formulated in terms of least-squares estimation. For m neighboring points $\mathbf{p}_1, \ldots, \mathbf{p}_m$, let $\mathbf{u}_i = (u_i^1, u_i^2, u_i^3)^t = Q(\mathbf{p}_i - \mathbf{p})$. Then the unknown coefficients β_i are chosen to be the least-squares estimates of a system of linear equations $Y = \mathbb{X}\beta$: where $\beta = (\beta_1, \cdots, \beta_5)^t$, $Y = (u_1^3, \cdots, u_m^3)^t$ and the $m \times 5$ matrix $\mathbb{X}$ is given by

$$\mathbb{X} = \begin{pmatrix} u_1^1 & u_1^2 & (u_1^1)^2 & u_1^1 u_1^2 & (u_1^2)^2 \\ u_2^1 & u_2^2 & (u_2^1)^2 & u_2^1 u_2^2 & (u_2^2)^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ u_m^1 & u_m^2 & (u_m^1)^2 & u_m^1 u_m^2 & (u_m^2)^2 \end{pmatrix}.$$

The least-squares estimation is

$$\hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_5)^t = (\mathbb{X}^t\mathbb{X})^-\mathbb{X}^t Y, \tag{3.26}$$

where $^-$ denotes a generalized inverse, which can be obtained through the singular value decomposition [68]. Then from (3.24), the fundamental forms can be estimated and consequently the mean and the Gaussian curvatures. In practice, g_{ii} can be any number bigger than 1. The smoother the triangular mesh, the closer g_{11}, g_{22} are to the value 1. When $g_{ij} = \delta_{ij}$, we have locally Euclidean space.

3.3.4 Local Conformal Coordinates

Because $g_{ij} \neq \delta_{ij}$, the Laplace-Beltrami operator (5.1) has many terms. As we have seen in the previous section, the mathematical form (5.1) reduces to the planar Laplacian when locally conformal coordinates are used. Such local conformal coordinate can be obtained by a linear transformation. Let us define new coordinates $v = (v^1, v^2)^t = A^{-1}u$, where $A = (a_{ij})$ is an invertible constant matrix in the neighborhood of $\mathbf{p} = X(u)$. Then we have the new local surface parameterization $Y(v) \equiv X(u) = X(Av)$. Let $g^* = (g^*_{ij})$ be the new metric matrix corresponding to the new parameterization. Then from the chain rule, $\frac{\partial Y}{\partial v^i} = \sum_k a_{ki} \frac{\partial X}{\partial u^k}$ and

$$g^*_{ij} = \langle \frac{\partial Y}{\partial v^i}, \frac{\partial Y}{\partial v^j} \rangle = \sum_{k,l} a_{ki} a_{lj} \langle \frac{\partial X}{\partial u^k}, \frac{\partial X}{\partial u^l} \rangle = \sum_{k,l} a_{ki} a_{lj} g_{kl}.$$

In matrix notation, $g^* = A' g A$. Since g is symmetric positive definite, there exists $g^{-1/2} = (g^{-1/2}_{ij})$. So by choosing $A = g^{-1/2}$ in the neighborhood of $\mathbf{p}$, $g^*_{ij} = \delta_{ij}$ and we have the local conformal coordinates (v^1, v^2). In such a coordinate system, the surface (3.25) can be written as

$$z(v^1, v^2) = \gamma_1 (v^1)^2 + \gamma_2 v^1 v^2 + \gamma_3 (v^2)^2 + \cdots$$

for some constants $\gamma_1, \gamma_2, \cdots$ and the Laplace-Beltrami operator becomes

$$\Delta_X = \frac{\partial^2}{\partial (v^1)^2} + \frac{\partial^2}{\partial (v^2)^2}.$$

3.3.5 Least-Squares Estimation of the Laplace-Beltrami operator

Using the above conformal mapping, we have reduced the problem of estimating the Laplace-Beltrami on an irregular triangulation in 3D to the problem of estimating the planar Laplacian on an irregular triangulation in a 2D flat plane. Estimating the planar Laplacian on an irregular triangulation can be solved as a least-squares estimation problem [4, 56].

The i-th neighboring point $\mathbf{p}_i$ of $\mathbf{p}$ under the local conformal transformation becomes $\mathbf{v}_i = (v_i^1, v_i^2, v_i^3)^t = g^{1/2}Q(\mathbf{p}_i - \mathbf{p})$. Now expand the function $F = F(v)$ as a Taylor expansion and evaluate at $\mathbf{v}_i$:

$$F(\mathbf{v}_i) \approx F(0) + F_1(0)v_i^1 + F_2(0)v_i^2 + \frac{1}{2}F_{11}(0)(v_i^1)^2 + F_{12}(0)v_i^1 v_i^2 + \frac{1}{2}F_{22}(0)(v_i^2)^2.,$$

where $F_j = \frac{\partial F}{\partial v^j}$ and $F_{ij} = \frac{\partial^2 F}{\partial v^i \partial v^j}$. Let

$$\alpha = (\alpha_1, \ldots, \alpha_5)^t = \big(F_1(0), F_2(0), F_{11}(0), F_{12}(0), F_{22}(0)\big)^t$$

and

$$Y = \big(F(\mathbf{v}_1) - F(0), \ldots, F(\mathbf{v}_m) - F(0)\big)^t.$$

Then we solve a system of linear equations $Y = \mathbb{X}\alpha$, where the i-th row of the matrix $\mathbb{X}$ is

$$\Big(v_i^1, v_i^2, \frac{1}{2}(v_i^1)^2, v_i^1 v_i^2, \frac{1}{2}(v_i^2)^2\Big).$$

The least-squares estimation to $Y = \mathbb{X}\alpha$ is $\hat{\alpha} = (\hat{\alpha}_1, \ldots, \hat{\alpha}_5) = (\mathbb{X}^t\mathbb{X})^-\mathbb{X}^t Y$, where $^-$ denotes a generalized inverse. Therefore, the Laplace-Beltrami operator of F at vertex $\mathbf{p}$ is estimated by

$$\widehat{\Delta F}(\mathbf{p}) = \hat{\alpha}_3 + \hat{\alpha}_5 = \sum_{i=1}^{m} w_i\big(F(\mathbf{p}_i) - F(\mathbf{p})\big),$$

where the weight w_i is equivalent to the the sum of the i-th component of the 3rd and 5th rows of $(\mathbb{X}'\mathbb{X})^-\mathbb{X}^t$. We will show in the next section how the weights w_i can be estimated via the finite element method. However, it has been shown that the approximation of the Laplacian based on least-squares estimation is one of the best performers on an irregular triangulation of a sphere [56]. The explanation lies in the fact that the above least-squares approximation is based on a Taylor expansion, thus explicitly minimizing the difference between the analytical and the approximated Laplacian.

3.4 Finite Element Method

In this section we present a finite element method (FEM) to solve the diffusion equation on the cortical surface. The FEM has its origin in the field of structural analysis and since then it has been used in diverse areas. Because the ASP algorithm already provides a triangular

mesh, there is no need to discretize the surface $\partial\Omega$ so it has the advantage of avoiding local surface parameterizations. In general, the finite element discretization of a PDE requires a corresponding variational principle [93]. However, in our approach we have avoided using the variational method.

3.4.1 Barycentric Coordinates

Let N_T be the number of triangles in the triangular mesh. We seek an approximate solution F_i within the i-th triangle T_i such that the solution $F_i(\mathbf{x}, t)$ is continuous across neighboring triangles. The approximate solution F for the whole region is then

$$F(\mathbf{x}, t) \approx \sum_{i=1}^{N_T} F_i(\mathbf{x}, t). \tag{3.27}$$

The most common form of approximation for F within each triangular element is linear interpolation [93]. Let $\mathbf{p}_{i_1}, \mathbf{p}_{i_2}, \mathbf{p}_{i_3}$ be the vertices of a triangular element T_i. We interpolate the solution F_i by

$$F_i(\mathbf{x}, t) = \xi_{i_1}(\mathbf{x})F(\mathbf{p}_{i_1}, t) + \xi_{i_2}(\mathbf{x})F(\mathbf{p}_{i_2}, t) + \xi_{i_3}(\mathbf{x})F(\mathbf{p}_{i_3}, t), \tag{3.28}$$

where ξ_{i_k} are given by the *barycentric coordinates,* which is a special case of *element shape functions* [99, 93]. It is possible to expand the solution F at each vertex of the triangulation, which has been used in estimating the Laplace-Beltrami operator in the brain surface flattening problem [6]; however, from (3.27) and (3.28), we immediately see that these two expansions are equivalent.

In the barycentric coordinates, any point $\mathbf{x} \in T_i$ is uniquely determined by

$$\mathbf{x} = \xi_{i_1}(\mathbf{x})\mathbf{p}_{i_1} + \xi_{i_2}(\mathbf{x})\mathbf{p}_{i_2} + \xi_{i_3}(\mathbf{x})\mathbf{p}_{i_3} \text{ and } \xi_{i_1}(\mathbf{x}) + \xi_{i_2}(\mathbf{x}) + \xi_{i_3}(\mathbf{x}) = 1. \tag{3.29}$$

If $\mathbf{x} \notin T_i$, we may let $\xi_{i_k} = 0$. So the barycentric coordinates satisfy $0 \le \xi_{i_k} \le 1$. Solving the equations (3.29), we get

$$(\xi_{i_1}, \xi_{i_2})^t = (P^t P)^{-1} P^t (\mathbf{x} - \mathbf{p}_{i_3}),$$

where the 3×2 matrix $P = (\mathbf{p}_{i_1} - \mathbf{p}_{i_3}, \mathbf{p}_{i_2} - \mathbf{p}_{i_3})$.

3.4.2 Discrete Diffusion Equation

Let G be an arbitrary piecewise linear function given by

$$G(\mathbf{x}) = \sum_{i=1}^{N_T} \xi_{i_1}(\mathbf{x})G_{i_1} + \xi_{i_2}(\mathbf{x})G_{i_2} + \xi_{i_3}(\mathbf{x})G_{i_3},$$

where $G_{i_1}, G_{i_2}, G_{i_3}$ are the values of function G at the three vertices of the triangle T_i. Note that the Laplace-Beltrami operator is self-adjoint with respect to the L^2 norm defined on the space of continuous piecewise linear functions [51]. Thus the diffusion equation (4.24) at each triangular element T_i becomes

$$\int_{T_i} G\frac{\partial F}{\partial t}\, dS = \int_{T_i} G\Delta F\, dS = -\int_{T_i} \langle \nabla F, \nabla G\rangle\, dS. \tag{3.30}$$

The left-hand term in (3.30) can be written as

$$\int_{T_i} G\frac{\partial F}{\partial t}\, dS = \sum_{k,l=1}^{3} G_{i_k}\frac{\partial F(\mathbf{p}_{i_l}, t)}{\partial t}\int_{T_i} \xi_{i_k}\xi_{i_l}\, dS = [G_i]^t[A^i]\frac{d}{dt}[F_i], \tag{3.31}$$

where $[G_i]^t = (G_{i_1}, G_{i_2}, G_{i_3})$, $[F_i] = \big(F(\mathbf{p}_{i_1}, t), F(\mathbf{p}_{i_2}, t), F(\mathbf{p}_{i_3}, t)\big)^t$ and the kl-th element of the matrix $[A^i]$ is $A^i_{kl} = \int_{T_i} \xi_{i_k}\xi_{i_l}\, dS$. It can be shown that

$$[A^i] = \frac{|T_i|}{12}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix},$$

where $|T_i|$ is the area of the triangular element T_i [93, pp. 459-465]. Similarly the right-hand term in (3.30) can be written as

$$\int_{T_i} \langle \nabla F, \nabla G\rangle\, dS = \sum_{k,l=1}^{3} G_{i_k}F(\mathbf{p}_{i_l}, t)\int_{T_i} \langle \nabla\xi_{i_k}, \nabla\xi_{i_l}\rangle\, dS = [G_i]^t[C^i][F_i], \tag{3.32}$$

where the kl-th element of $[C^i]$ is $C^i_{kl} = \int_{T_i}\langle \nabla\xi_{i_k}, \nabla\xi_{i_l}\rangle\, dS$. Because T_i is planar, the gradient ∇ becomes the standard 2-dimensional Euclidean gradient. The matrix $[C^i]$ is usually called the *element coefficient matrix* and its exact expression can be computed using the property of the element shape functions [99, 93]:

$$[C^i] = \frac{1}{2}\begin{pmatrix} \cot\theta_{i_2} + \cot\theta_{i_3} & -\cot\theta_{i_3} & -\cot\theta_{i_2} \\ -\cot\theta_{i_3} & \cot\theta_{i_1} + \cot\theta_{i_3} & -\cot\theta_{i_1} \\ -\cot\theta_{i_2} & -\cot\theta_{i_1} & \cot\theta_{i_1} + \cot\theta_{i_2} \end{pmatrix},$$

where θ_{i_k} is the interior angle of vertex $\mathbf{p}_{i_k}$ of the triangle T_i. From (3.31) and (3.32), the equation (3.30) becomes

$$[G_i]^t[A^i]\frac{d[F_i]}{dt} = -[G_i]^t[G^i][F_i]. \tag{3.33}$$

Since the equation (3.33) should be satisfied for an arbitrary column vector $[G_i]^t$, we have a system of ordinary differential equation given by

$$\frac{d[F_i]}{dt} = -[A^i]^{-1}[C^i][F_i] \tag{3.34}$$

for each triangle T_i. The equation (3.34) is a discretized diffusion equation within each triangle T_i.

3.4.3 Assembling Elements

Having discretized a triangular element, the next step is to assemble all such elements in m incident triangles around the central node $\mathbf{p}$. We do not need to assemble all elements in the triangular mesh but only m incident triangles around the node $\mathbf{p}$ because the diffusion of signal for a relatively small time interval is strictly a local phenomenon. Let $\mathbf{p}_1, \cdots, \mathbf{p}_m$ be the m neighboring nodes around $\mathbf{p} = \mathbf{p}_0$ in the counter-clockwise direction. Let $\mathbf{p}, \mathbf{p}_i, \mathbf{p}_{i+1}$ be the vertices of the triangular element T_i (Figure 3.5). Then from (3.32),

$$\begin{aligned}
\int_{T_1 \cup \cdots \cup T_m} \langle \nabla F, \nabla G \rangle \, dS &= \sum_{i=1}^{m} \int_{T_i} \langle \nabla F, \nabla G \rangle \, dS && (3.35)\\
&= \sum_{i=1}^{m} [G_i]^t[C^i][F_i] && (3.36)\\
&= [G]^t[C][F], && (3.37)
\end{aligned}$$

where the column vectors are

$$[F] = [F(\mathbf{p}, t), F(\mathbf{p}_1, t), \cdots, F(\mathbf{p}_m, t)]^t,$$

$$[G] = [G(\mathbf{p}), G(\mathbf{p}_1), \cdots, G(\mathbf{p}_m)]^t.$$

The $(m+1)\times(m+1)$ matrix $[C]$ is called the *global coefficient matrix*, which is the assemblage of individual element coefficients. The contribution to the ij-th element C_{ij} of the matrix

$[C]$ comes from all finite elements containing nodes i and j $(0 \le i, j \le m)$. In the case of a hexagonal triangulation in Figure 3.5, the 7×7 matrix $[C]$ is given by

$$\begin{pmatrix} C_{00}^1 + \cdots + C_{00}^6 & C_{01}^1 + C_{01}^6 & C_{02}^1 + C_{02}^2 & C_{03}^2 + C_{03}^3 & C_{04}^3 + C_{04}^4 & C_{05}^4 + C_{05}^5 & C_{06}^5 + C_{06}^6 \\ C_{01}^1 + C_{01}^6 & C_{11}^1 + C_{11}^6 & C_{12}^1 & 0 & 0 & 0 & C_{16}^6 \\ C_{02}^1 + C_{02}^2 & C_{12}^1 & C_{22}^1 + C_{22}^2 & C_{23}^2 & 0 & 0 & 0 \\ C_{03}^2 + C_{03}^3 & 0 & C_{23}^2 & C_{33}^2 + C_{33}^3 & C_{34}^3 & 0 & 0 \\ C_{04}^3 + C_{04}^4 & 0 & 0 & C_{34}^3 & C_{44}^3 + C_{44}^4 & C_{45}^4 & 0 \\ C_{05}^4 + C_{05}^5 & 0 & 0 & 0 & C_{45}^4 & C_{55}^4 + C_{55}^5 & C_{56}^5 \\ C_{06}^5 + C_{06}^6 & C_{16}^6 & 0 & 0 & 0 & C_{56}^5 & C_{66}^5 + C_{66}^6 \end{pmatrix}.$$

Also from (3.31),

$$\int_{T_1 \cup \cdots \cup T_m} G \frac{\partial F}{\partial t}\, dS = \sum_{i=1}^{m} \int_{T_i} G \frac{\partial F}{\partial t}\, dS \tag{3.38}$$

$$= \sum_{i=1}^{m} [G_i]^t [A^i] \frac{d[F_i]}{dt} \tag{3.39}$$

$$= [G]^t [A] \frac{d[F]}{dt}, \tag{3.40}$$

where the ij-th element A_{ij} of the matrix $[A]$ has the same structure as $[C]$, i.e. $A_{01} = A_{01}^1 + A_{01}^6$ instead of $C_{01} = C_{01}^1 + C_{01}^6$ in the first row and the second column. Combining (3.37) and (3.40), we have

$$[G]^t [A] \frac{d[F]}{dt} = -[G]^t [C][F]. \tag{3.41}$$

Since equation (3.41) should be satisfied for an arbitrary piecewise linear function G, we have a discrete diffusion equation on m elements $T_1, \cdots, T_m$ given by

$$\frac{d[F]}{dt} = -[A]^{-1}[C][F]. \tag{3.42}$$

The first row of the simultaneous ODE (3.42) gives the discrete diffusion equation at the vertex $\mathbf{p} = \mathbf{p}_0$:

$$\frac{dF(\mathbf{p}, t)}{dt} = -\sum_{i,k=0}^{m} A_{0k}^{-1} C_{ki} F(\mathbf{p}_i, t), \tag{3.43}$$

where A_{0k}^{-1} is the $0k$-th element of A^{-1}. The right-hand side of the equation (3.43) is the estimation of the Laplace-Beltrami operator based on the FEM at $\mathbf{p}$. Simplifying the matrix

computation using the computational algebra software Maple, we have the FEM estimation for the Laplace-Beltrami operator given by

$$\widehat{\Delta F}(\mathbf{p}) = \sum_{i=1}^{m} w_i \big(F(\mathbf{p}_i) - F(\mathbf{p})\big) \tag{3.44}$$

with the weights

$$w_i = \frac{\cot\theta_i + \cot\phi_i}{|T|},$$

where θ_i and ϕ_i are the two angles opposite to the edge $\mathbf{p}_i - \mathbf{p}$ and $|T| = \sum_{i=1}^{m} |T_i|$ is the sum of the areas of the incident triangles (Figure 3.5). In terms of Cartesian coordinates, we have

$$\cot\theta_i = \frac{\langle \mathbf{p}_{i+1} - \mathbf{p}, \mathbf{p}_{i+1} - \mathbf{p}_i \rangle}{2|T_i|}, \cot\phi_i = \frac{\langle \mathbf{p}_{i-1} - \mathbf{p}, \mathbf{p}_{i-1} - \mathbf{p}_i \rangle}{2|T_i|} \text{ and}$$

$$|T_i| = \frac{1}{2}\|(\mathbf{p}_{i+1} - p) \times (\mathbf{p}_i - p)\|.$$

When $m = 4$ with the fixed angles $\theta_i, \phi_i = \pi/4$, the triangular mesh becomes a square grid and $\widehat{\Delta F}(\mathbf{p})$ in (5.1) reduces to the finite difference estimation of the Laplacian in the square grid [56, 102].

3.5 Finite Difference Scheme

So far we have presented two methods of estimating the Laplace-Beltrami operator as a weighted linear smoothing of the form $\widehat{\Delta F}(\mathbf{p}, t) = \sum_{i=1}^{m} w_i \big(F(\mathbf{p}_i, t) - F(\mathbf{p}, t)\big)$. In both methods, the diffusion equation can be solved by the finite difference scheme:

$$F(\mathbf{p}, t_{n+1}) = F(\mathbf{p}, t_n) + (t_{n+1} - t_n)\widehat{\Delta F}(\mathbf{p}, t_n) \tag{3.45}$$

with the initial condition $F(\mathbf{p}, 0) = f(\mathbf{p})$ for each node $\mathbf{p}$ on the triangular mesh. We may fix the iteration step size $t_{n+1} - t_n = \delta t$. The value δt controls the spread of the diffusion smoothing and it should be chosen with respect to the smoothness of the Laplace-Beltrami operator of F. If $\mathbf{p}_1, \cdots, \mathbf{p}_m$ are neighboring nodes of $\mathbf{p}$, then the diffusion of heat should satisfy the following approximate boundary condition for small δt

$$\min_{0 \le i \le m} F(\mathbf{p}_i, t_n) \le F(\mathbf{p}, t_{n+1}) \le \max_{0 \le i \le m} F(\mathbf{p}_i, t_n). \tag{3.46}$$

The inequality (3.46) simply states that the diffused heat must be anywhere between the highest and the lowest heat. Since the Laplace-Beltrami operator can be transformed to the the planar Laplacian via local conformal transform, we only need to show (3.46) is in $\mathbb{R}^2$. Let $A_{\mathbf{p}} \subset \mathbb{R}^2$ be a bounded set containing a point $\mathbf{p}$. Suppose that

$$\max_{\mathbf{x} \in A_{\mathbf{p}}} F(\mathbf{x}, t) = F(\mathbf{p}_{max}, t).$$

For infinitesimally small δt, $F(\mathbf{p}, t + \delta t)$ restricted to $A_{\mathbf{p}}$ will also attain it's maximum at $\mathbf{p}_{max}$. At the local maximum $\mathbf{p}_{max}$, the second derivatives of F will be negative. Hence $\Delta F(\mathbf{p}_{max}, t) \leq 0$. Then

$$\begin{aligned} F(\mathbf{p}, t + \delta t) &\leq F(\mathbf{p}_{max}, t + \delta t) \\ &= F(\mathbf{p}_{max}, t) + \int_t^{\delta t} \Delta F(\mathbf{p}_{max}, s)\, ds \\ &\leq \max_{\mathbf{x} \in A_{\mathbf{p}}} F(\mathbf{p}, t). \end{aligned}$$

The minimum case follows similarly. The inequality (3.46) may break down if δt is large. From (3.45) and (3.46), the iteration step size must satisfy

$$\delta t \leq \min\Big(\Big| \frac{\max_i F(\mathbf{p}_i, t_n) - F(\mathbf{p}, t_n)}{\widehat{\Delta F}(\mathbf{p}, t_n)} \Big|, \Big| \frac{\min_i F(\mathbf{p}_i, t_n) - F(\mathbf{p}, t_n)}{\widehat{\Delta F}(\mathbf{p}, t_n)} \Big| \Big). \tag{3.47}$$

The denominator $\widehat{\Delta F}$ behaves like the sample covariance of F at $\mathbf{p}, \mathbf{p}_1, \cdots, \mathbf{p}_m$. The smoother the function F is, the smaller the Laplace-Beltrami operator of F is. In such a case, the iteration step size δt can be large. By changing the iteration step size δt with respect to the inequality at each node p, we will have *spatially adaptive smoothing*, which depends on the smoothness of the function F. In order to determine the critical iteration step size for the isotropic diffusion (with respect to the local conformal coordinates) with the spatially fixed δt, it is best to measure the smoothness of the function F first and then estimate the δt accordingly. The signal $F(\mathbf{p}, t_n)$ tends to become smoother as n increases, in which case, the ratio in (3.47) gets larger. Therefore, if δt satisfies the inequality (3.47) at every node $\mathbf{p}$ at the first iteration $n = 1$, the inequality (3.46) will be satisfied for the later iteration $n \geq 1$. Note that the iteration time step should be in the order of

$$\delta t < \text{nodal distance} \times \Big| \frac{\text{the first derivative of F}}{\text{the second derivatives of F}} \Big|$$

to guarantee the convergence and the stability of the finite difference scheme. The relation between the critical iteration step size and the nodal distance has also been briefly pointed out in [4].

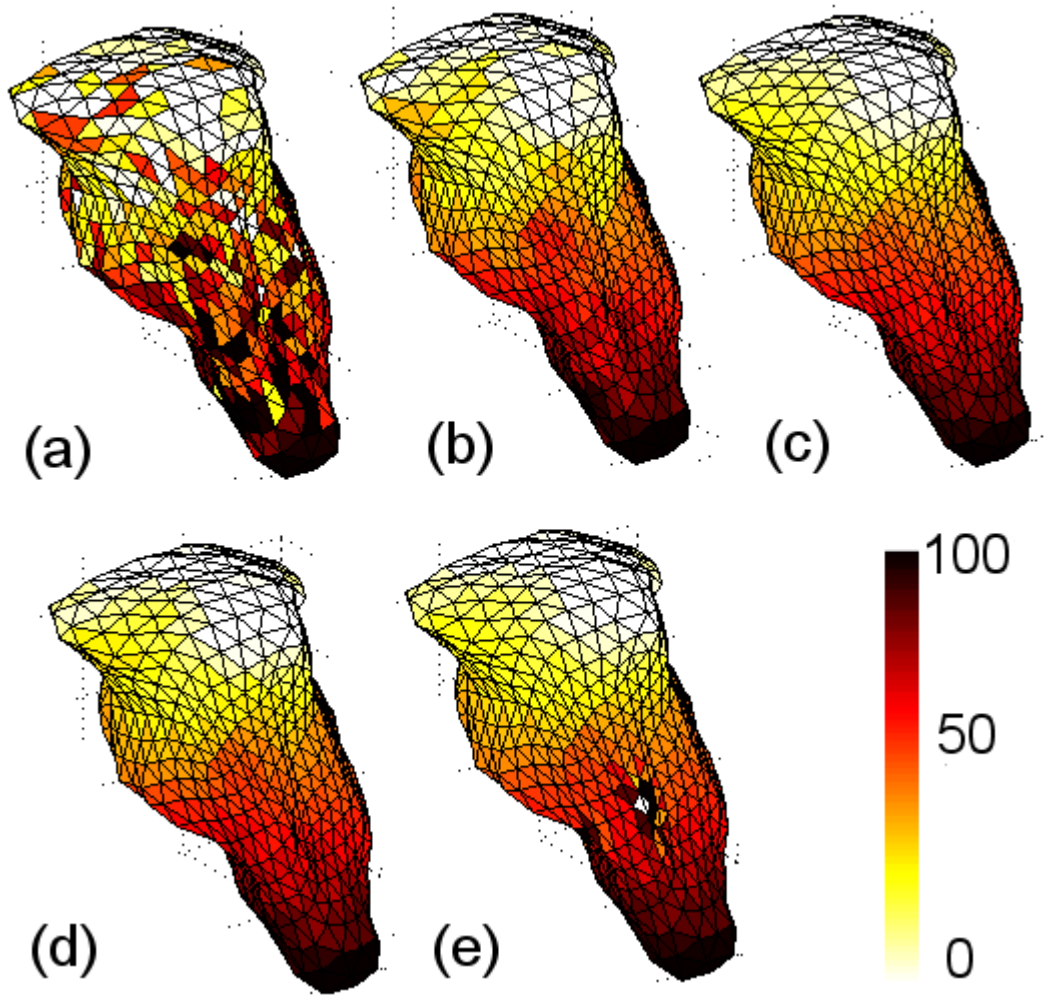

Figure 3.6: Diffusion smoothing of an artificial heat distribution on the triangulated mesh of the brain stem consisting of 1280 triangles. The artificial signal was generated with Gaussian noise to illustrate how the finite difference scheme works with different iteration step sizes. **a.** The initial heat distribution, **b.** After 10 iterations with $\delta t = 0.5$, **c.** After 20 iterations with $\delta t = 0.5$, **d.** After 50 iterations with $\delta t = 0.2$, **e.** After 10 iterations with $\delta t = 1.5$. Because the iteration step size is large, the iteration breaks down and a singularity (the white spot) begins to appear.

If the iteration step size is bigger than the desired inequality, the iteration will diverge as illustrated in Figure 3.6 **e.**, where the large iteration step size $\delta t = 1.5$ causes the instability of the iterations producing the sudden singularity (the white spot). Ideally, the finite difference scheme should converge to the stationary solution of the diffusion equation, i.e. $\Delta F = 0$. After N iterations, the finite difference scheme gives an approximate solution of the diffusion of the initial heat f after time $N\delta t$. If the diffusion were applied to Euclidean space, it would be equivalent to the Gaussian Kernel smoothing with FWHM $= 4(\ln 2)^{1/2}\sqrt{N\delta t}$. In order to have 10mm FWHM Gaussian kernel smoothing in Euclidean space, we should have $N\delta t = 4.33$. Hence if the iteration step size is taken as $\delta t = 0.2$, then $N = 22$ iterations are sufficient to get the 10mm FWHM smoothing assuming the iterations are stable. Equivalently $\delta t = 0.1$ with $N = 44$ will have the equivalent result. So the number of iterations that is required is inversely proportional to the iteration step size δt. The smaller the iteration step size, the longer it takes to achieve the same result. This has been illustrated in Figure 3.6 **c.** and **d.** which show almost identical results. It should be remembered that the above discussion about 10mm FWHM Gaussian kernel smoothing is only an analogy applied to the curved surface and should not be taken literally. To see this, note that the Laplace-Beltrami operator in general orthogonal coordinate system (u^1, u^2) can be written as

$$\Delta = \frac{\partial^2}{\partial (u^1)^2} + \frac{\partial^2}{\partial (u^2)^2} + \rho_2 \frac{\partial}{\partial u^1} + \rho_1 \frac{\partial}{\partial u^2},$$

where ρ_1, ρ_2 are the *geodesic curvatures* of the u^1, u^2 axes [64]. In the conformal coordinate system we are using, ρ_1, ρ_2 are zeros. So locally in the conformal coordinates, the FWHM has the same meaning as in Euclidean space. But in general with respect to the orthogonal coordinates, the Laplace-Beltrami operator involves the geodesic curvatures of the surface. Therefore, how the surface is curved influences the smoothing and FWHM does not have the same meaning as in Euclidean space.

3.6 Result: Smoothing the Mean Curvature

The surface based diffusion smoothing has been used to increase the signal to noise ratio in fMRI activation on the cortical surface [4, 48]. Another application of the smoothing technique would be in the area of enhancing structural information such as the mean and the

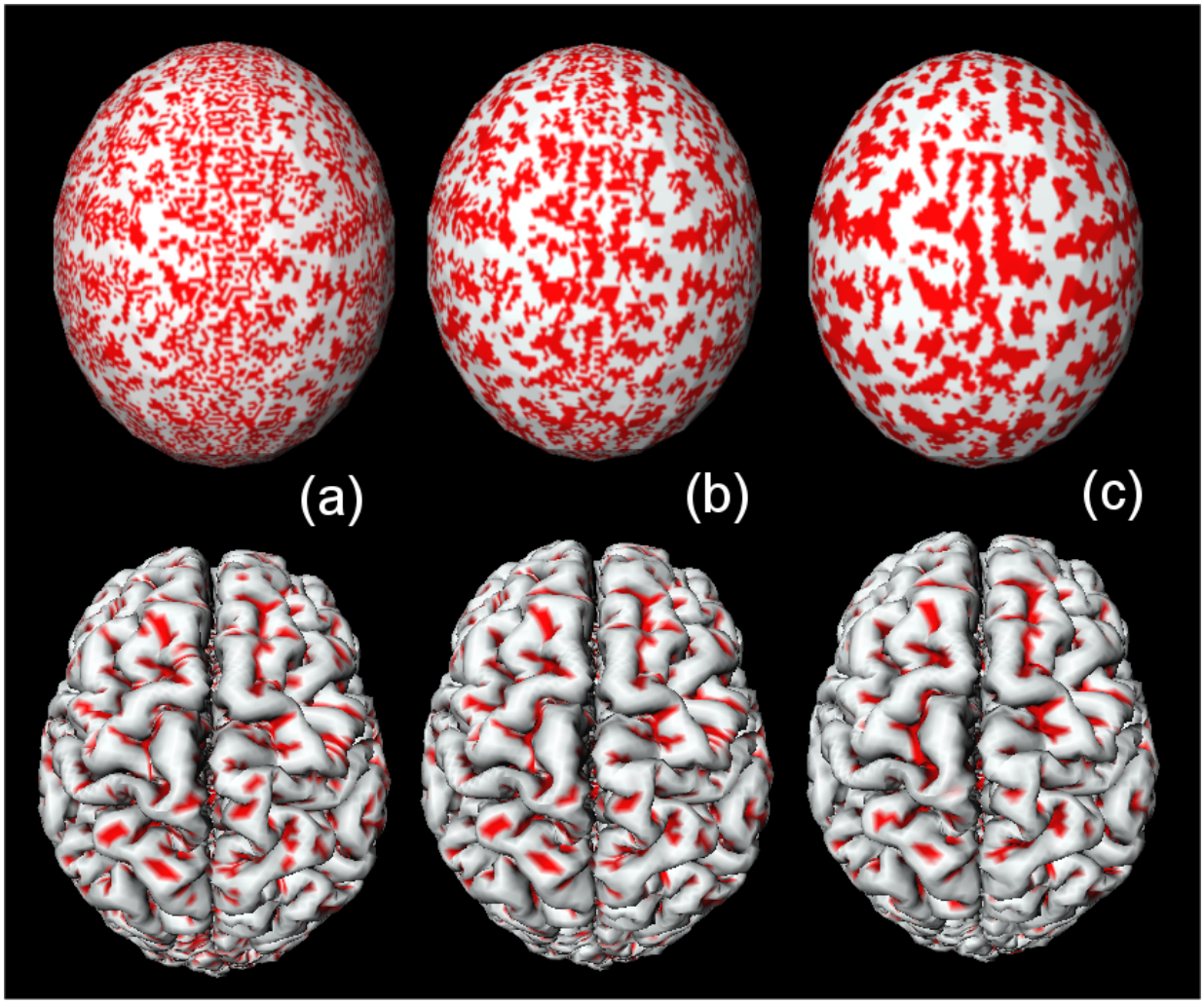

Figure 3.7: The mean curvature of the outer cortex is mapped onto an ellipsoid consisting of 81920 triangles preserving the connectivity. Note that the diffusion was run directly on the cortical surface and mapped onto the ellipsoid later. **a.** Before the iteration. **b.** After 40 iterations with $\delta t = 0.02$ **c.** After 100 iterations with $\delta t = 0.02$. If the smoothing were based on simple internodal averaging, such sulcal pattern can not be obtained.

Gaussian curvatures on the outer cortex [60], the cortical thickness [58], and the displacement vector fields on the cortical surface deformation problem. Just like smoothing functional activation, it is also possible to smooth coordinates of the triangular mesh resulting in a mesh smoothing problem [78]. So with slight modification of our algorithm, triangular mesh smoothing might be another application of the diffusion smoothing although care should be taken to avoid the inherent mesh shrinkage problem [105].

Among many possible applications related to brain surfaces, we have picked up an example of segmenting sulci and gyri of the cortical surface based on the segmentation of the mean curvature. Sulci and gyri can be characterized as the crowns and the hollows of the

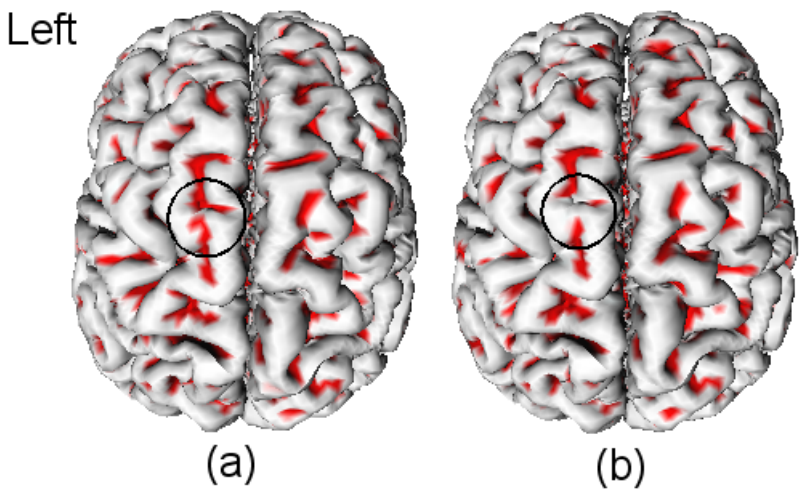

Figure 3.8: Comparison of the parametric method and the finite element method of the diffusion smoothing of the mean curvature in Figure 3.7 **c.** The extent of smoothing is roughly equivalent to 5mm FWHM Gaussian kernel smoothing in 2D Euclidian space ($\delta t = 0.02$, $N = 100$ iterations). **a.** Parametric method. **b.** The finite element method. The smoothing patterns are slightly different inside the black circle.

brain surface [103]. Gaussian and mean curvatures of the brain surface have been used to characterize its shape [33, 50, 60]. Then this geometric information can be further used as landmarks for the brain registration process. The problem of segmenting sulci and gyri has usually been done via multiscale or multiresolution methods in computing curvatures. However, one can achieve similar result using the diffusion smoothing. As shown in Figure 3.7 **a.**, **b.** and **c.**, the maximum mean curvature can identify the sulci although they also identify some unwanted regions which do not belong to the sulci. Then by applying diffusion smoothing to the mean curvature of the cortical surface, the sulcal regions can be enhanced (Figure 3.7 **d**, **e.** and **f.**). Basically the same principle has been used in extracting the edge [79, 42].

There are two ways to compute the curvatures. The FEM version [37, 78] estimates the mean curvature K_M as

$$K_M = \frac{1}{4|T|} \sum_{i=1}^{m} (\cot \theta_i + \cot \phi_i) \langle \mathbf{n}, \mathbf{p}_i - \mathbf{p} \rangle.$$

In the parametric version, which we will be using in our example (Figures 3.7), the mean curvature K_M is estimated using the formula (3.17). Then in terms of the local quadratic surface (3.25), we have the mean curvature estimation (3.19).

Comparing the parametric method and the finite element method of smoothing the mean curvature, we have similar but nonidentical diffusion profiles due to, perhaps, different assumptions on the manifold structure where the Laplace-Beltrami operators were actually computed (Figure 3.8). But the major reason for the slight discrepancy is that 81920 triangles are not sufficient enough to guarantee the smoothness of the triangulation, which is essential for both methods of estimation to converge.

Chapter 4

Deformation-Based Morphometry on 2D Surfaces

This chapter is the extension of the Chapter 2, where we have introduced a unified statistical approach for *deformation-based morphometry* in 3D MRI. Using the same stochastic assumption on the deformation field as before, we derive the statistical distributions of the morphological variables such as area dilatation rate, cortical thickness and curvature changes. These statistics can be used in statistical inferences on surface-based morphometric analysis. As an illustration, we will demonstrate how the surface-based statistical analysis can be applied in localizing the cortical regions of gray matter tissue growth and loss in the brain images longitudinally collected in the same group of children and adolescents previously analyzed in Chapter 2.

4.1 Surface Deformation

Let $U(x,t) = (U_1, U_2, U_3)^t$ be the 3-dimensional displacement vector required to deform the structure at position $x = (x_1, x_2, x_3)$ in the gray matter Ω_0 to the homologous structure after time t. The whole volume Ω_0 will deform continuously and smoothly to Ω_t via the deformation $x \to x + U$ while the cortical boundary $\partial\Omega_0$ will deform to $\partial\Omega_t$. Note that the cortical surface $\partial\Omega_t$ consists of two parts: the outer cortical surface $\partial\Omega_t^{out}$ between the gray

matter and CSF and the inner cortical surface $\partial\Omega_t^{in}$ between the gray and white matter (see Figure 1.1 and Figure 4.1), i.e.

$$\partial\Omega_t = \partial\Omega_t^{out} \cup \partial\Omega_t^{in}.$$

Although we are dealing exclusively with the cortical boundaries, our surface-based analysis can be equally applicable to the deformation of the surface of any brain substructure.

Any statistical inference on structural deformation requires a basic stochastic model. The proposed stochastic model on the displacement velocity V, which has been already used in Chapter 2 is

$$V(x) = \mu(x) + \Sigma^{1/2}(x)\epsilon(x), x \in \Omega_0, \tag{4.1}$$

where μ is the mean displacement velocity and $\Sigma^{1/2}$ is the 3×3 symmetric positive definite covariance matrix, which allows for correlations between components of the displacement fields. The components of the error vector ϵ are are assumed to be independent and identically distributed as smooth stationary Gaussian random fields with zero mean and unit variance. The statistical model (4.1) is based on the whole gray matter volume Ω_0 so it is not truly a surface based model. However by restricting the domain of the displacement U to the cortical surface $\partial\Omega_0$, we have the surface based model:

$$V(x) = \mu(x) + \Sigma^{1/2}(x)\epsilon(x), x \in \partial\Omega_0. \tag{4.2}$$

It can be shown that the normal component of the displacement velocity $V = \frac{\partial U}{\partial t}$ restricted on the boundary $\partial\Omega_0$ uniquely determine the evolution of the cortical surface. Assuming the surface $\partial\Omega_t$ to be smooth enough, it can be *locally* expressed in an implicit form

$$F(x,t) = 0, x \in \partial\Omega_t \tag{4.3}$$

By taking the time derivative in (4.3), the *kinematic equation* [39, pp. 33] for the surface deformation is given by

$$\frac{\partial F}{\partial t} + \langle V, \nabla F \rangle = 0, \tag{4.4}$$

where $\nabla F = \left(\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \frac{\partial F}{\partial x_3}\right)^t$ is the gradient vector. The unit normal vector to the surface is given by

$$\mathbf{n} = \frac{\nabla F}{\|\nabla F\|}. \tag{4.5}$$

From (4.4) and (4.5), the kinematic equation becomes

$$\frac{\partial F}{\partial t} = -\|\nabla F\| V_{\mathbf{n}}, \tag{4.6}$$

where $V_{\mathbf{n}} = \langle V, \mathbf{n} \rangle$ is the normal component of the surface displacement velocity. If we let $V_{\mathbf{t}}$ denote the tangential component of V, then $V = V_{\mathbf{n}} + V_{\mathbf{t}}$. There are infinitely many surface displacement velocities that gives the same normal component $V_{\mathbf{n}}$ and in turn, the same kinematic equation (4.6), which describes the evolution of the cortical surface over time. Hence, translation of the surface in the tangential direction does not change the geometry of the surface and only the normal component $V_{\mathbf{n}}$ uniquely determines the evolution of the cortical surface at a given point.

The major impediment to the practical use of cortical surface-based approaches in brain imaging is the difficulty of automating the surface extraction, registration and analysis. Cortical surfaces are usually extracted as triangular meshes via the Marching Cubes algorithm [72], the level set method [97] or deformable surfaces method [33]. In our analysis, we have used the anatomic segmentation using proximities (ASP) method [74], which is a variant of the deformable surfaces method. In triangulating cortical surfaces of the human brain from 3D MRI, the ASP method generates 81,920 triangles evenly distributed in size. In order to accomplish the statistical analysis on the cortical surface, mathematical representation of the cortical surface is an essential part. The most natural mathematical representation of the cortical surface is by *surface parameterization* [14, 18, 64, 76]. We model the cortical surface as a smooth 2-dimensional Riemannian manifold parameterized by two parameters u^1 and u^2 such that any point $x \in \partial\Omega_0$ can be uniquely represented by

$$x = X(u^1, u^2)$$

for some parameter space $u = (u^1, u^2) \in D \subset \mathbb{R}^2$. We will try to avoid global parameterization such as tensor B-splines in (3.22), which are computationally expensive compared to a local surface parameterization. Instead, a quadratic polynomial

$$z(u^1, u^2) = \beta_1 u^1 + \beta_2 u^2 + \beta_3 (u^1)^2 + \beta_4 u^1 u^2 + \beta_5 (u^2)^2$$

will be used as a local parameterization fitted via least-squares estimation. Using the least-squares solution in (3.26), these coefficients β_i can be estimated. There is a slightly different quadratic surface parameterization, which has been used in estimating curvatures of a macaque monkey brain surface [60].

4.2 Surface-Based Morphological descriptors

4.2.1 Riemannian Metric Tensor Change

As in the case of local volume change in the whole brain volume, the rate of cortical surface area expansion or reduction may not be uniform across the cortical surface. Extending the concept of volume dilatation, we introduce a new concept of surface area dilatation and its rate of change over time via differential geometry.

Suppose that the cortical surface $\partial\Omega_t$ at time t can be parameterized by the parameters $u = (u^1, u^2)$ such that any point $x \in \partial\Omega_t$ can be written as $x = X(u,t)$. Following the convention of differential geometry, we will suppress the spatial parameter u in $X(u,t)$ and write it as $X(t)$ whenever there is no ambiguity. Then we have

$$X(t) = X(0) + U(X(0), t). \tag{4.7}$$

Let $X_i = \frac{\partial X}{\partial u^i}$ be a partial derivative vector defined in (3.14). The *Riemannian metric tensor* g_{ij} of the surface $\partial\Omega_t$ is given by $g_{ij}(t) = \langle X_i(t), X_j(t)\rangle$. The Riemannian metric tensor g_{ij} measures the amount of the deviation of the cortical surface from a flat Euclidean plane. It enables us to measure lengths, angles and areas in the cortical surface and that is why g_{ij} is called the *metric tensor*. For a flat plane, $g_{ij} = \delta_{ij}$. Differentiating (4.7) with respect to the parameter u^i,

$$X_i(t) = X_i(0) + (\nabla U)X_i(0), \tag{4.8}$$

where $\nabla U = (\frac{\partial U_l}{\partial x_k})$ is the 3×3 displacement gradient matrix defined in (2.4). The metric tensor $g_{ij}(t)$ can be written as

$$\begin{aligned} g_{ij}(t) &= \langle X_i(t), X_j(t)\rangle & (4.9)\\ &= g_{ij}(0) + 2X_i^t(0)(\nabla U)X_j(0) + X_i^t(0)(\nabla U)^t(\nabla U)X_j(0), & (4.10) \end{aligned}$$

where ^{t} is the matrix transpose. For relatively small displacement, the higher order term involving $(\nabla U)^t(\nabla U)$ can be neglected:

$$g_{ij}(t) \approx g_{ij}(0) + 2X_i(0)^t(\nabla U)X_j(0).$$

Let $g(t) = \big(g_{ij}(t)\big)$ be a 2×2 matrix of metric tensors of $\partial\Omega_t$. In matrix notation, the rate of metric structure change is given by

$$\frac{\partial g}{\partial t} \approx 2(\nabla X)^t(\nabla V)\nabla X, \tag{4.11}$$

where $\nabla X = \big(X_1(0), X_2(0)\big)$ is a 3×2 gradient matrix.

4.2.2 Local Surface Area Change

The *infinitesimal surface area element* [64, pp. 114] is defined as

$$\sqrt{\det g} = (g_{11}g_{22} - g_{12}^2)^{1/2}. \tag{4.12}$$

It measures the transformed area of the unit square in the parameter space D via the transformation $X : D \to \partial\Omega_t$. The total surface area of $\partial\Omega_t$ is then given by

$$\|\partial\Omega_t\| = \int_{X^{-1}(\partial\Omega_t)} \sqrt{\det g}\, du^1 du^2.$$

Thus, the *local area dilatation rate* Λ_{area} or the rate of local surface area change per unit surface area becomes

$$\Lambda_{area} = \frac{\partial}{\partial t} \ln \sqrt{\det g} = \frac{1}{2 \det g} \frac{\partial(\det g)}{\partial t}.$$

If the whole brain volume Ω_t is parameterized by 3-dimensional curvilinear coordinates $u = (u^1, u^2, u^3)$, then $\frac{\partial}{\partial t} \ln \sqrt{\det g}$ is equivalent to the local volume dilatation rate Λ_{volume}, discussed in (2.5). Therefore, in terms of the curvilinear coordinate system, the area dilatation and volume dilatation are the same concept. A simple matrix manipulation [54, pp. 304-308] shows that

$$\Lambda_{area} = \frac{1}{2}\mathrm{tr}\Big(g^{-1}\frac{\partial g}{\partial t}\Big). \tag{4.13}$$

From (4.11) and (4.13), the rate of local surface area change becomes

$$\Lambda_{area} \approx \mathrm{tr}[g^{-1}(\nabla X)^t(\frac{\partial}{\partial t}\nabla U)\nabla X].$$

The partial derivatives of Gaussian random fields are again Gaussian [1, pp. 31]. Under the stochastic model (4.1), the area dilatation rate is distributed as Gaussian:

$$\Lambda_{area} = \lambda_{area}(x) + \epsilon_{area}(x), \tag{4.14}$$

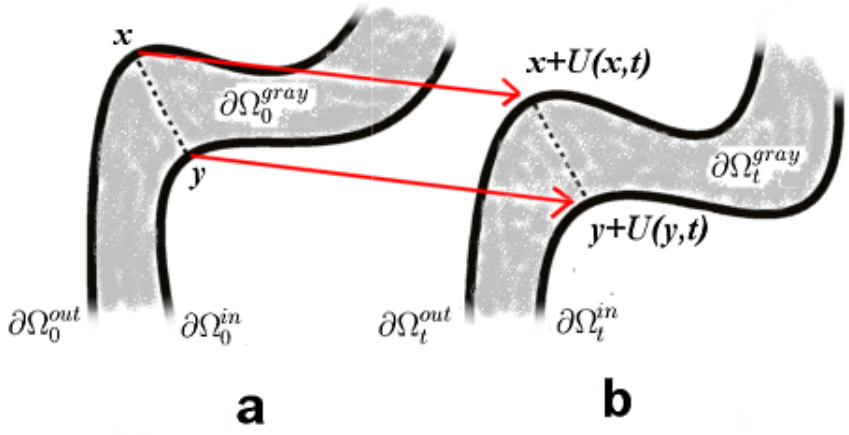

Figure 4.1: Cortical thickness change under the deformation $x \to x + U(x,t)$. The linkage between x and y is defined by the ASP algorithm and the cortical thickness is defined as the Euclidian distance $\|x - y\|$. **a.** Before the deformation. **b.** After the deformation.

where $\lambda_{area}(x) = tr[g^{-1}(\nabla X)^t(\nabla\mu)\nabla X]$ is the mean area dilatation rate and $\epsilon_{area}(x)$ is a mean zero Gaussian random field defined on the cortical surface. The area dilatation rate is invariant under parameterization, i.e. the area dilatation rate will always be the same no matter which paramerization is chosen. Afterwards, statistical inference of brain tissue growth near the cortical surface can be performed via the T random field defined on the cortical surface [120, 125]. As in the case of local volume dilatation model (2.6), the $\mathbf{Var}(\epsilon_{area})$ should not depend on t in order to apply the random field theory developed in [120, 125].

4.2.3 Cortical Thickness Change

The average cortical thickness for each individual is about 3mm [55]. Cortical thickness varies from 2mm to 4mm depending on the location of the cortex. When the brain develops, it is highly likely that the change of cortical thickness may not be uniform across the cortical surface. We will show how to localize the cortical regions of statistically significant thickness change in brain development. Our approach introduced here can also be applied to measuring the rate of cortical thinning, possibly associated with Alzheimer's disease. As in the case of volume dilatation, we introduce the concept of length dilatation, which measures cortical thickness change per unit thickness. There are many different computational approaches to measuring cortical thickness but we will use the Euclidean distance from a point x on the

outer surface $\partial\Omega_t^{out}$ to the corresponding point y on the inner surface $\partial\Omega_t^{in}$, as defined by the automatic linkages used in the ASP algorithm [74] (Figure 4.1). A validation study for the assessment of the accuracy of the cortical thickness measure based on the ASP algorithm has been performed and found to be valid for the most of the cortex [61]. There is an alternate method for automatically measuring cortical thickness based on the Laplace equation [58].

Let $\|x-y\|$ be the cortical thickness computed as the Euclidean distance between $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$, i.e.

$$\|x-y\| = \big((y_1-x_1)^2 + (y_2-x_2)^2 + (y_3-x_3)^2\big)^{1/2}.$$

Under the deformation (4.7), the cortical thickness at $x \in \partial\Omega_t^{out}$ can be written as

$$\|x(t)-y(t)\| = \|x(0)-y(0)+U(x(0),t)-U(y(0),t)\|$$

For relatively small displacement, we may neglect the higher order terms of U in the Taylor expansion of the cortical thickness:

$$\|x(t)-y(t)\| \approx \|x(0)-y(0)\| + \big(U^t(x(0),t)-U^t(y(0),t)\big)\frac{x(0)-y(0)}{\|x(0)-y(0)\|}.$$

Furthermore, $U(x(0),t)-U(y(0),t) \approx \nabla U(x(0),t)\big(x(0)-y(0)\big)$. It follows that

$$\frac{\partial}{\partial t}\|x-y\| = \big(x(0)-y(0)\big)(\nabla V)\frac{x(0)-y(0)}{\|x(0)-y(0)\|}.$$

Hence the length dilatation rate Λ_{length} is given as a quadratic form in $\mathbf{d}$ such that

$$\Lambda_{length} = \frac{\partial}{\partial t}\ln\|x-y\| \approx \mathbf{d}^t(\nabla V)\mathbf{d} = \sum_{i,j=1}^{3} d_i d_j \frac{\partial^2 U_j}{\partial t \partial x_i},$$

where the unit vector $\mathbf{d} = (d_1, d_2, d_3)^t = \frac{x(0)-y(0)}{\|x(0)-y(0)\|}$. From (4.1), we have a linear model on the length dilatation given by

$$\Lambda_{length} = \lambda_{length}(x) + \epsilon_{length}(x),$$

where $\lambda_{length} = \mathbf{d}^t\big(\nabla\mu\big)\mathbf{d}$ is the mean cortical thickness dilatation rate and ϵ_{length} is a mean zero Gaussian random field. In practice, the cortical thickness dilatation rate Λ_{length}^j for subject j is given by the discrete approximation:

$$\Lambda_{length}^j = \frac{\|x(t_j)-y(t_j)\| - \|x(0)-y(0)\|}{t_j\|x(0)-y(0)\|},$$

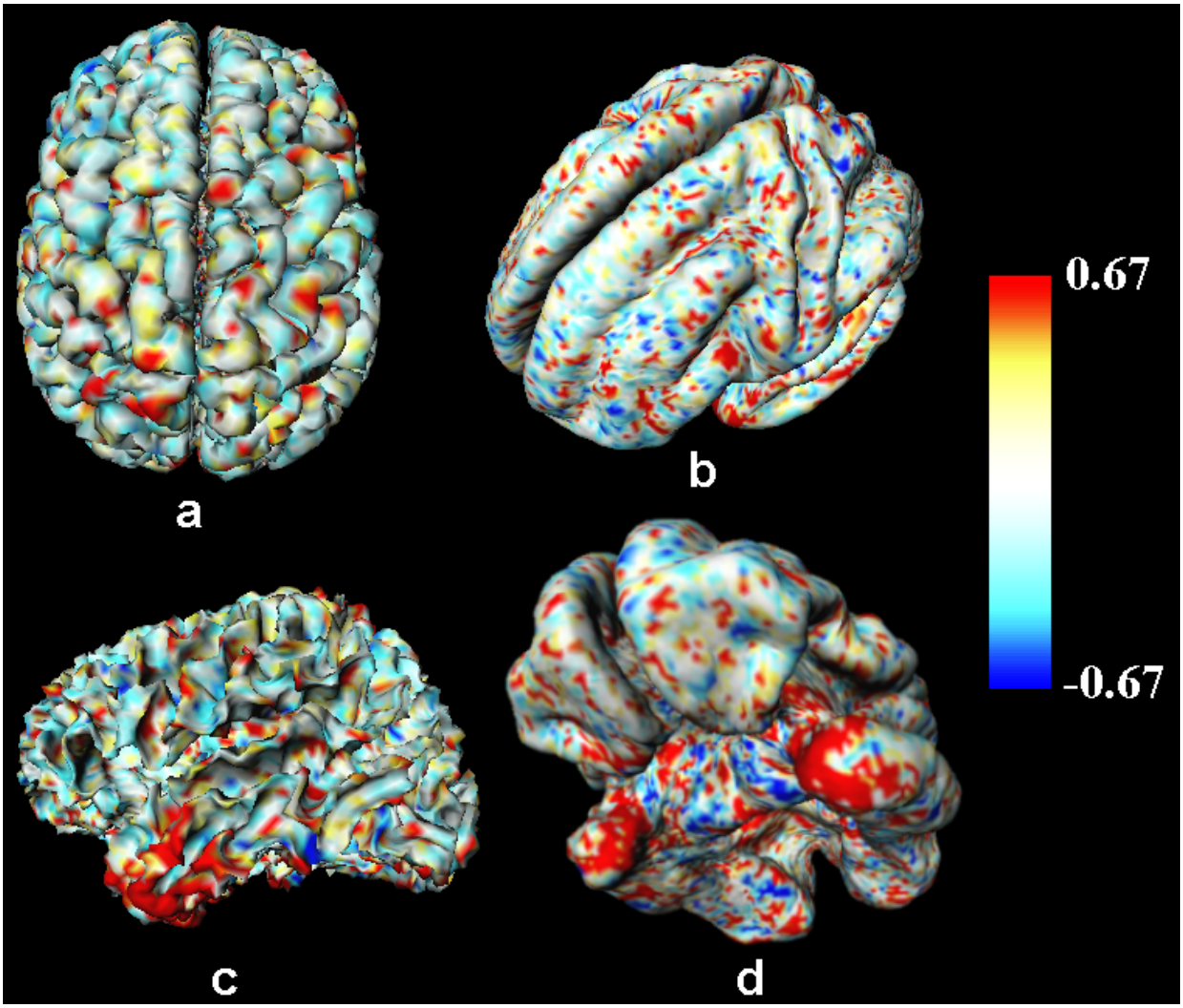

Figure 4.2: Cortical thickness dilatation rate for a single subject. The red (blue) regions show more than 67% thickness increase (decrease). **a.** The outer cortical surface. **b., d.** The same data as in **a.** projected onto the average outer cortical surface of 28 subjects. **c.** The inner cortical surface. Note the high variabilities of the cortical thickness dilatation rate across the cortex. Due to such large variabilities, surface-based smoothing is required to increase the signal-to-noise ratio.

where t_j is the time difference between two scans. Afterwards the t-statistic is formed by

$$T_{length} = \sqrt{n}\frac{M_{length}}{S_{length}},$$

where M_{length} and S_{length} are the sample mean and standard deviation of n subjects respectively.

4.2.4 Curvature Change

When the surface $\partial\Omega_0$ deforms to $\partial\Omega_t$, curvatures of the surface change. The mean and the Gaussian curvature can characterize the shape and location of the sulci and gyri, which are the crowns and hollows of the brain surface [60, 103]. By measuring the rate of change of the mean and Gaussian curvature, rapidly folding and unfolding cortical regions can be localized. Let us first consider a special case of the displacement restricted to $U(x,t) = l(t)\mathbf{n}(x,t)$, where l is independent of $x \in \partial\Omega_0$. Under this parallel deformation, the deformed surface $\partial\Omega_t$ is called a *parallel surface* of $\partial\Omega_0$ and its parametric form $x(u,t)$ is given by

$$X(u,t) = X(u,0) + l(t)\mathbf{n}(u,t).$$

Then the mean curvature $K_M(x,t)$ and Gaussian curvature $K_G(x,t)$ at $x \in \partial\Omega_t$ can be computed (see [108, pp. 102-107]) as

$$K_M(x,t) = \frac{K_M(x,0) - 2K_G(x,0)l(t)}{1 + K_G(x,0)l^2(t) - 2K_M(x,0)} \tag{4.15}$$

$$K_G(x,t) = \frac{K_G(x,0)}{1 + K_G(x,0)l^2(t) - 2K_M(x,0)l(t)} \tag{4.16}$$

For relatively small displacement, the first two terms of Taylor expansions in (4.15) and (4.16) are given by

$$K_M(x,t) \approx K_M(x,0) + 2\big(K_M^2(x,0) - K_G(x,0)\big)l(t), \tag{4.17}$$

$$K_G(x,t) \approx K_G(x,0) + 2K_M(x,0)K_G(x,0)l(t). \tag{4.18}$$

Then from (4.17) and (4.18), the rate of the curvature changes over time are

$$\frac{\partial K_M}{\partial t} = 2\big(K_M^2(0) - K_G(0)\big)\frac{dl}{dt}$$

$$\frac{\partial K_G}{\partial t} = 2K_M(0)K_G(0)\frac{dl}{dt}.$$

In general, the displacement vector field U is not always parallel to the surface normal $\mathbf{n}$ in the deformation of brain and we need to generalize the concept of parallel surface in differential geometry to nonparallel surface. Based on the kinematic equation (4.4), the rate of the change of curvatures are given as a system of simultaneous partial differential equations [39, pp. 206-210]. Let κ_1 and κ_2 be the two principal curvatures defined in Section 3.2.2. then it can be shown that the rate of the curvature change are

$$\frac{\partial \kappa_i}{\partial t} = \kappa_i^2 V_{\mathbf{n}} + \Delta V_{\mathbf{n}}, \tag{4.19}$$

where Δ is the Laplace-Beltrami operator defined on the surface. From (3.17), which relates the principal curvatures to the mean and Gaussian curvatures, we have

$$\begin{aligned} \frac{\partial K_G}{\partial t} &= 2K_M K_G V_{\mathbf{n}} + 2K_M \Delta V_{\mathbf{n}} \\ \frac{\partial K_M}{\partial t} &= 2(K_M^2 - K_G)V_{\mathbf{n}} + \frac{1}{2}\Delta V_{\mathbf{n}}. \end{aligned}$$

For relatively small displacement, we can neglect the higher order derivatives of $V_{\mathbf{n}}$,

$$\frac{\partial K_G}{\partial t} \approx 2K_M K_G V_{\mathbf{n}}, \tag{4.20}$$

$$\frac{\partial K_M}{\partial t} \approx 2(K_M^2 - K_G)V_{\mathbf{n}}. \tag{4.21}$$

From the statistical model (4.1), the normal velocity component is

$$V_{\mathbf{n}} = \langle V, \mathbf{n} \rangle = \mu_{\mathbf{n}} + \epsilon_V, \tag{4.22}$$

where $\mu_{\mathbf{n}} = \langle \mu, \mathbf{n} \rangle$ is the mean normal velocity and ϵ_V is a mean zero Gaussian random field. It follows that the mean curvature change is,

$$\frac{\partial K_M}{\partial t} = 2(K_M^2 - K_G)\mu_{\mathbf{n}} + \epsilon_{K_M},$$

where ϵ_{K_M} is a mean zero Gaussian random field.

4.3 Diffusion Smoothing

In order to increase the signal-to-noise ratio, kernel smoothing is desirable in many statistical analyses. For example, Figure 4.2 shows high noise data on the outer cortical surface of

the average brain atlas. By smoothing the data on the curved surface, the signal-to-noise ratio will increase if the signal itself is smooth and in turn, it will be easier to localize the morphological changes. However, due to the convoluted geometry of the cortex, which is non-Euclidean, we can not directly apply Gaussian kernel smoothing on the cortical surface. Gaussian kernel smoothing of the data $f(x), x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ with FWHM (*full width at half maximum*) $= 4(\ln 2)^{1/2}\sqrt{t}$ is defined as the convolution of the Gaussian kernel with f:

$$F(x,t) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-(x-y)^2/4t} f(y) dy. \tag{4.23}$$

As we have introduced earlier in Chapter 3, the convoluted data $F(x,t)$ is the integral solution of the n-dimensional diffusion equation

$$\frac{\partial F}{\partial t} = \Delta F \tag{4.24}$$

with the initial condition $F(x,0) = f(x)$, where $\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$ is the Laplacian in n-dimensional Euclidean space [41]. Hence the Gaussian kernel smoothing of the function $f(x)$ is equivalent to the diffusion of the initial heat $f(x)$ after time t. The indirect approach of solving the PDE (4.24) rather than Gaussian kernel smoothing gives *diffusion smoothing*, which is adaptable to curved surfaces by generalizing the Laplacian Δ defined in $\mathbb{R}^n$ to Riemannian manifolds. When using diffusion smoothing on curved surfaces, the smoothing somehow has to incorporate the geometrical features of the curved surface and the Laplacian Δ should change accordingly. The extension of the Euclidean Laplacian to an arbitrary Riemannian manifold is called the *Laplace-Beltrami operator* [64]. Diffusion smoothing on the cortical surface has been used in the problem of the smoothing fMRI to increase the signal-to-noise ratio (SNR) [4] (See [38, 91, 124] for the precise definition of the SNR). The approach taken in [4] is a local flattening of the cortical surface and estimating the planar Laplacian, which may not be as accurate as our estimation based on the parametric and the finite element method. There should be further comparative investigation between these two methods. However, instead of flattening the cortical surface first and then doing Gaussian kernel smoothing, it is possible to solve the diffusion equation on the curved surface via the Laplace-Beltrami operator. If the surface $\partial\Omega$ is given by the parameterization $X = X(u^1, u^2)$, the surface Laplacian Δ should take a different form which is determined by the Riemannian metric tensor g_{ij}. From [7, pp. 158-167], we can show that the Laplace-Beltrami operator

takes the following metric tensor formulation:

$$\Delta F = \sum_{i,j} \frac{1}{|g|^{1/2}} \frac{\partial}{\partial u^i} \Big(|g|^{1/2} g^{ij} \frac{\partial F}{\partial u^j} \Big), \tag{4.25}$$

where $(g^{ij}) = g^{-1}$. Such a surface Laplacian is called the *Laplace-Beltrami* operator corresponding to the parameterization X. Using the FEM on the triangular cortical mesh generated by the ASP algorithm, we estimated the Laplace-Beltrami operator as the linear weights of neighboring data in Chapter 3. Let $\mathbf{p}_1, \cdots, \mathbf{p}_m$ be m neighboring nodes around the central node $\mathbf{p} = \mathbf{p}_0$. Then we have shown in (5.1) that the estimated Laplace-Beltrami operator is

$$\widehat{\Delta F}(p) = \sum_{i=1}^{m} w_i \big(F(\mathbf{p}_i) - F(\mathbf{p}) \big)$$

with the weights

$$w_i = \frac{\cot \theta_i + \cot \phi_i}{\sum_{i=1}^{m} \|T_i\|},$$

where θ_i and ϕ_i are the two angles opposite to the edge connecting $\mathbf{p}_i$ and $\mathbf{p}$, and $\|T_i\|$ is the area of ith triangle (Figure 3.5). Note that this is an improved formulation from the previous attempt of diffusion smoothing used in smoothing fMRI data in [4], where the Laplacian is estimated as the planar Laplacian after local fattening of the triangular mesh consisting of nodes $\mathbf{p}_0, \cdots, \mathbf{p}_m$ onto a flat plane. In the actual numerical implementation using Matlab, we have used formulas

$$\cot \theta_i = \frac{\langle \mathbf{p}_{i+1} - \mathbf{p}, \mathbf{p}_{i+1} - \mathbf{p}_i \rangle}{2\|T_i\|}, \cot \phi_i = \frac{\langle \mathbf{p}_{i-1} - \mathbf{p}, \mathbf{p}_{i-1} - \mathbf{p}_i \rangle}{2\|T_i\|}$$

and $\|T_i\| = \frac{1}{2}\|(\mathbf{p}_{i+1} - \mathbf{p}) \times (\mathbf{p}_i - \mathbf{p})\|$. Afterwards, the finite difference scheme is used to iteratively solve the diffusion equation at each node $\mathbf{p}$:

$$\frac{F(\mathbf{p}, t_{n+1}) - F(\mathbf{p}, t_n)}{t_{n+1} - t_n} = \sigma^2 \widehat{\Delta} F(\mathbf{p}, t_n),$$

with the initial condition $F(\mathbf{p}, 0) = f(\mathbf{p})$. After N iterations, the finite difference scheme gives an approximate solution of the diffusion of the initial heat f after time $N\delta t$. If the diffusion were applied to Euclidean space, it would be equivalent to Gaussian kernel smoothing with

$$\text{FWHM} = 4(\ln 2)^{1/2} \sqrt{N \delta t}.$$

It should be emphasized that Gaussian kernel smoothing is a special case of diffusion smoothing restricted to Euclidian space. Computing the linear weights for the Laplace-Beltrami

operator takes a fair amount of time (the parametric method takes about 3 minutes and FEM method takes about 6 minutes in Matlab running on a Pentium III machine), but once the weights are computed, it is applied through the whole iteration repeatedly and the actual finite difference scheme takes only two minutes for 100 iterations.

4.4 Statistical Inference on 2D Surfaces

All of our morphological descriptors $\Lambda(x)$ such as surface area dilatation, cortical thickness change, curvature change are modeled as Gaussian random fields on the respective cortical surfaces, i.e.

$$\Lambda(x) = \lambda(x) + \epsilon(x), x \in \partial\Omega_{atlas}, \tag{4.26}$$

where the deterministic part λ is the mean of the morphological descriptor Λ and ϵ is a mean zero Gaussian random field. As we have explained earlier in (2.6), we need to assume that $\mathbf{Var}(\epsilon)$ does not depends on time t. In order to do statistical inference about structural changes, we need to map these morphological descriptors to a template cortical surface $\partial\Omega_{atlas}$. The T random field on the manifold $\partial\Omega_{atlas}$ is defined as

$$T(x) = \sqrt{n}\frac{M(x)}{S(x)},\ x \in \partial\Omega_{atlas}$$

where M and S are the sample mean and standard deviation of the morphological descriptor Λ over the n subjects. Under the null hypothesis

$$H_0 : \lambda(x) = 0 \text{ for all } x \in \partial\Omega_{atlas},$$

i.e. no structural change based on the morphological descriptor Λ, $T(x)$ becomes a student's t-distribution with $n-1$ degrees of freedom at each fixed voxel x. The p-value of the local maxima of the T field will give a conservative threshold, which has been used in brain imaging for a quite some time now [125]. From (2.9), we have

$$P\Big(\max_{x\in\partial\Omega_{atlas}} T(x) \geq y\Big) \approx \sum_{i=0}^{3} \phi_i(\partial\Omega_{atlas})\rho_i(y), \tag{4.27}$$

where the Minkowski functional ϕ_i are

$$\phi_0(\partial\Omega_{atlas}) = 2, \phi_1(\partial\Omega_{atlas}) = 0, \phi_2(\partial\Omega_{atlas}) = \|\partial\Omega_{atlas}\|, \phi_3(\partial\Omega_{atlas}) = 0$$

and $\|\partial\Omega_{atlas}\|$ is the total surface area of $\partial\Omega_{atlas}$. In order to have more smoothed signals, we apply diffusion smoothing with given FWHM to the morphological descriptor on Ω_{atals}. When diffusion smoothing with given FWHM is applied to $\Lambda(x), x \in \partial\Omega_{atlas}$, the 2-dimensional EC-density [125] becomes

$$\rho_2(y) = \frac{1}{\text{FWHM}^2}\frac{4\ln 2}{(2\pi)^{3/2}}\frac{\Gamma(\frac{n}{2})}{(\frac{n-1}{2})^{1/2}\Gamma(\frac{n-1}{2})} y\Big(1+\frac{y^2}{n-1}\Big)^{-(n-2)/2}.$$

Hence, the excursion probability on the cortical surface can be approximated by the following simple formula:

$$P\Big(\max_{x\in\partial\Omega_{atlas}} T(x) \geq y\Big) \approx 2P\big(T(x)\geq y\big) + \|\partial\Omega_{atlas}\|\rho_2(y).$$

The total surface area varies from subject to subject. We can approximate the total surface area $\|\partial\Omega_{atlas}\|$ by summing the area of each triangle in a triangulated surface. The total surface area of the average atlas brain is 275,800 mm^2, which is roughly the area of 53cm $\times$ 53cm sheet. Note that the surface area of the mean atlas brain is different from the mean surface area of 28 subjects (Table 4.1). When 20mm FWHM diffusion smoothing is used on the template surface $\partial\Omega_{atlas}$, 2.5% thresholding gives

$$\begin{aligned} P\Big(\max_{x\in\partial\Omega_{atlas}} T(x) &\geq 5.1\Big) \approx 0.025, \\ P\Big(\max_{x\in\partial\Omega_{atlas}} T(x) &\leq -5.1\Big) \approx 0.025. \end{aligned}$$

4.5 Detecting Global Surface Deformation

So far our analysis has concentrated on detecting local regions of rapid morphological changes on the cortical surface. Global morphological analysis is relatively easier than local analysis in terms of its modeling and numerical computation. As we have shown in Chapter 2, global volume dilatation rate can be approximately modeled as a Gaussian random variable. Similarly, global surface area dilation rate can be modeled with a Gaussian random variable. Global morphological measures are important in the characterization of brain deformation. The total area of the cortical surface $\partial\Omega_t$ is given by

$$\|\partial\Omega_t\| = \int_{X^{-1}(\partial\Omega_t)} \sqrt{\det(g)}\, du,$$

where g is the metric matrix corresponding to the global parameterization $X(u)$ of the surface $\partial\Omega_t$. $\|\partial\Omega_t\|$ can be estimated by the sum of the areas of 81,920 triangles generated by the ASP algorithm. Then we define the total surface area dilatation as

$$\Theta_{total\ area} = \frac{\|\partial\Omega_t\| - \|\partial\Omega_0\|}{\|\partial\Omega_0\|}$$

and the total surface area dilatation rate as

$$\Lambda_{total\ area} = \frac{\partial}{\partial t} \ln \|\partial\Omega_t\| \Big|_{t=0}.$$

These measures will be used in showing that the total surface area decreases in both outer and inner cortical surfaces between ages 12 and 16. Note that

$$\Lambda_{total\ area} = \frac{1}{\|\partial\Omega_0\|} \int_D \frac{\partial}{\partial t} \sqrt{\det g}\, du \tag{4.28}$$

$$= \frac{1}{\|\partial\Omega_0\|} \int_D \Lambda_{area} \sqrt{\det g}\, du, \tag{4.29}$$

where $D = X^{-1}(\partial\Omega_0)$ and Λ_{area} is local surface area dilatation, which is distributed as Gaussian in (4.14). Hence, the total surface area dilatation is distributed as a Gaussian random variable and a statistical inference will be based on a simple t-test.

Let $h(x) = \|x - y\|$ be the cortical thickness measured at $x \in \partial\Omega_0^{out}$ with the linkage y defined by the ASP algorithm. The gray matter Ω_t can be considered as a thin shell bounded by two surfaces $\partial\Omega_t^{out}$ and $\partial\Omega_t^{in}$. Then the total volume of gray matter is approximately given by

$$\|\Omega_t\| \approx \int_{x \in \partial\Omega_0^{out}} h(x)\, dx \tag{4.30}$$

with respect to the outer cortical surface. (4.30) is only valid for relatively small thickness $h(x)$. A better approximation can be obtained by using the both outer and inner cortical surfaces at the same time. Since there are 81,920 triangles for each surfaces and each triangle has the corresponding triangle on the other surface, Ω_t consists of 81,920 triangular prisms. Let $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ be the three vertices of a triangle on the outer cortical surface and $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ be the corresponding three vertices on the inner cortical surface such that $\mathbf{p}_i$ is linked to $\mathbf{q}_i$ by ASP algorithm (Figure 4.3). The triangular prism consists of three tetrahedra with the vertices $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{q}_1\}$, $\{\mathbf{p}_2, \mathbf{p}_3, \mathbf{q}_1, \mathbf{q}_2\}$ and $\{\mathbf{p}_3, \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$. Then provided the sides of the

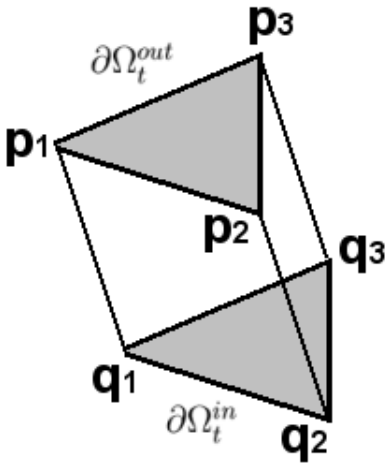

Figure 4.3: A triangular prism. 81,920 triangular prisms will form the gray matter.

prism are flat, the volumes of the triangular prism is given by

$$\begin{aligned} & \frac{1}{6}\Big|\det(\mathbf{p}_1-\mathbf{q}_1,\mathbf{p}_2-\mathbf{q}_1,\mathbf{p}_3-\mathbf{q}_1)\Big| \\ + \quad & \frac{1}{6}\Big|\det(\mathbf{p}_2-\mathbf{q}_2,\mathbf{p}_3-\mathbf{q}_2,\mathbf{q}_1-\mathbf{q}_2)\Big| \\ + \quad & \frac{1}{6}\Big|\det(\mathbf{q}_1-\mathbf{p}_3,\mathbf{q}_2-\mathbf{p}_3,\mathbf{q}_3-\mathbf{p}_3)\Big|. \end{aligned}$$

The total volume $\|\Omega_t\|$ can be estimated using the above discrete computation. Similarly we define the total gray-matter volume dilatation

$$\Theta_{gray} = \frac{\|\Omega_t\| - \|\Omega_0\|}{\|\Omega_0\|}$$

and the total gray-matter volume dilatation rate as

$$\Lambda_{gray} = \frac{\partial}{\partial t} \ln \|\Omega_t\| \Big|_{t=0}.$$

From (2.14),

$$\Lambda_{gray} \approx \frac{1}{\|\Omega_0\|} \int_{\Omega_0} \Lambda_{volume}\, dx,$$

where Λ_{volume} is local volume dilatation rate distributed as a mean zero Gaussian random field. So Λ_{gray} is approximately distributed as a mean zero Gaussian random variable.

4.6 Results

MR images of the same twenty eight normal subjects studied for the 3D deformation analysis in Chapter 2, were again used in the surface-based analysis. A triangular mesh for each

cortical surface was generated by deforming a mesh to fit the proper boundary in a classified volume using the ASP algorithm [74]. Each voxel was pre-classified as CSF, gray matter and white matter based on its intensity. For the first scan at time t_1, the outer cortical surface was triangulated in two steps: first, an ellipsoidal mesh placed outside the brain was shrunk down to the inner cortical surface, which is the white-gray matter boundary. The resulting mesh was used as the initial estimate in the second step that expands the mesh to fit the outer cortical surface, which is the gray-CSF boundary. To generate the outer surface for the second scan at time t_2, we start with the inner surface from the first scan taken at time t_1, and then expand it outwards to match the outer surface on the classified volume of the second scan. Starting with the same mesh for the inner surface in the two expansion steps, each node in the initial mesh gets mapped to a point on the outer surface for each scan. Since the two scans were affinely normalized to start with, two points to which the node gets mapped will be roughly homologous. From this modified ASP registration method, the displacement vector fields from the point on the outer surface of the first scan to the corresponding point on the second scan are obtained. The modified ASP registration method assumes heavily that the shape of the cortical surface does not appreciably change between the first scan and the second scan. This assumption is valid in the case of brain development for a short period of time as illustrated in Figure 1.2, where the global sulcal geometry remains stable for five year interval, although local cortical geometry shows some visible changes.

Total surface area dilatation. We measured the total surface area dilatation rate $\Lambda^j_{total\ area}$ for subject j by computing the total area of triangular meshes on the both outer and inner cortical surfaces. The mean total area dilatation rate was

$$\frac{1}{28}\sum_{j=1}^{28}\Lambda^j_{total\ area} = -0.0094.$$

This 0.9% decrease of the total cortical surface area per year is found to be statistically significant (t-value of -9.25). On average, there was 4.3% decrease in the total cortical surface area between the first scan taken at age 11.5 and the second scan taken at age 16.1. There has been substantial developmental studies on gray matter volume reduction for children and adolescents [30, 57, 80, 85, 89], but the ROI-based volumetry used in these studies did not allow investigators to detect the total surface area reduction in both outer

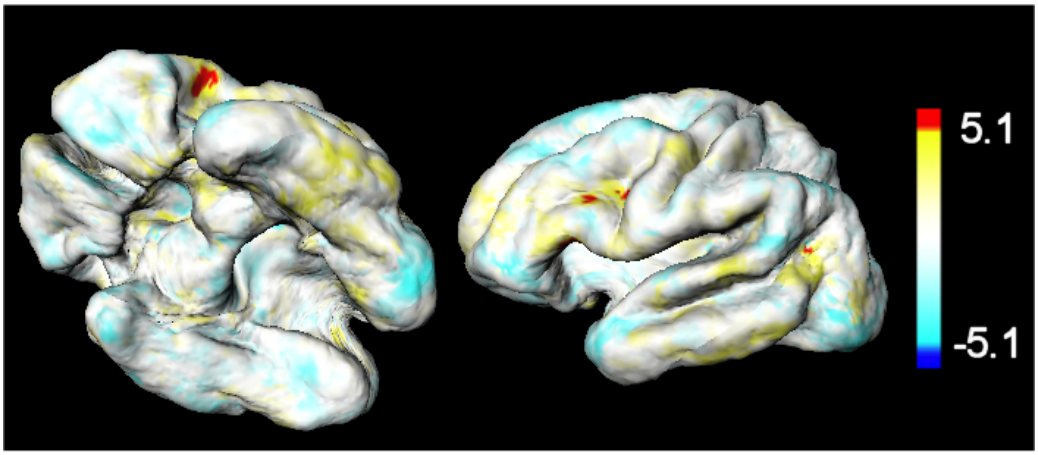

Figure 4.4: t-map of the cortical surface area dilatation rate showing the statistically significant region of area expansion and reduction. The red regions are statistically significant surface area expansions while the blue regions are statistically significant surface area reductions between ages 12 and 16. As in the case of local volume dilatation, it shows asymmetric growth patterns.

and inner cortical surfaces.

Local surface area dilatation. In order to localize surface area change, the surface area dilatation rates were computed for all subjects, then smoothed with 20mm FWHM diffusion smoothing to increase the signal-to-noise ratio. The growth pattern of cortical surface area change is different from that of the cortical thickness change. On average, local surface area changed from -15.79% to 13.78% per year. In one subject, we observed between -106.5% and 120.3% of the local surface area change over a 4 year time span. Figure 4.4 is the t-map of the cortical surface area dilatation. Surface area growth and decrease were detected by $T > 5.1$ and $T < -5.1$ ($P < 0.05$, corrected) respectively, showing statistically significant growth in localized temporal and parietal cortical regions of the left hemisphere and a localized area decrease in the right hemisphere. However, these relatively smaller regions of local surface area change may indicate that local surface area is not the dominant feature in brain development between ages 12 and 16.

Validation. To validate our surface-based morphometry, a small artificial bump was added to the triangular mesh of cerebellum with 1280 triangles with Gaussian random noise (Figure 4.5). Generating 30 such random cerebellum surfaces, we tried to see if local surface area change around the bump can be detected. By fitting the surfaces with local quadratic polynomials, we estimated the Riemannian metric tensor g_{ij} using (3.18). Local surface area

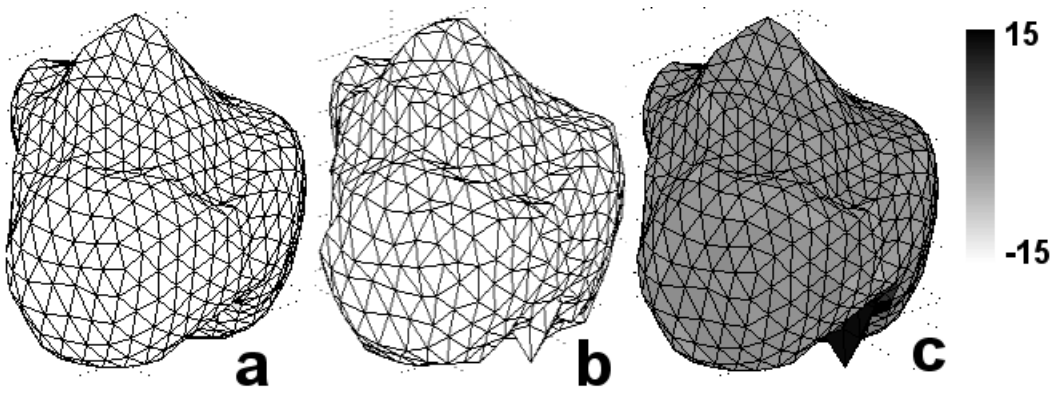

Figure 4.5: Simulation on artificial data. **a.** The cerebellum surface with 1280 triangles. **b.** 30 artificially deformed surfaces are generated with a bump at a fixed location with Gaussian noise. **c.** t-map of local surface area dilatation after 5mm FWHM diffusion.

element (4.12) is computed by

$$\sqrt{\det g} = (1 + \beta_1^2 + \beta_2^2)^{1/2}.$$

By subtracting the local surface area elements from **b** to **a** in Figure 4.5, we get surface dilatation for each random surface. Then 30 dilatation fields on the cerebellum surface are smoothed with 5mm FWHM diffusion smoothing and t-map is formed. **c** in Figure 4.5 shows very high t-value of 15 around the bump validating our methodology.

Total gray-matter volume dilatation. The total gray-matter volume dilatation rate Λ_{gray}^j for subject j was computed. The mean total gray-matter volume dilatation rate is

$$\frac{1}{28}\sum_{j=1}^{28} \Lambda_{gray}^j = -0.0050.$$

This 0.5% decrease in the gray-matter volume per year is statistically significant (t-value of -4.45). So we have the global morphological patterns of both shrinking cortical surface area and shrinking gray-matter volume between ages 12 and 16 although there is some localized cortical regions where cortical surface expansion has been detected.

Cortical thickness dilatation. Also we computed the mean cortical thickness dilatation rate $M_{thick}(x)$ at each voxel x defined by

$$M_{thick}(x) = \frac{1}{28}\sum_{j=1}^{28} \Lambda_{thick}^j(x).$$

Then the average cortical thickness dilatation across the cortical surface is given by

$$\frac{1}{\|\partial\Omega_{avg}\|}\int_{x\in\partial\Omega_{avg}} M_{thick}(x)\, dx = 0.026,$$

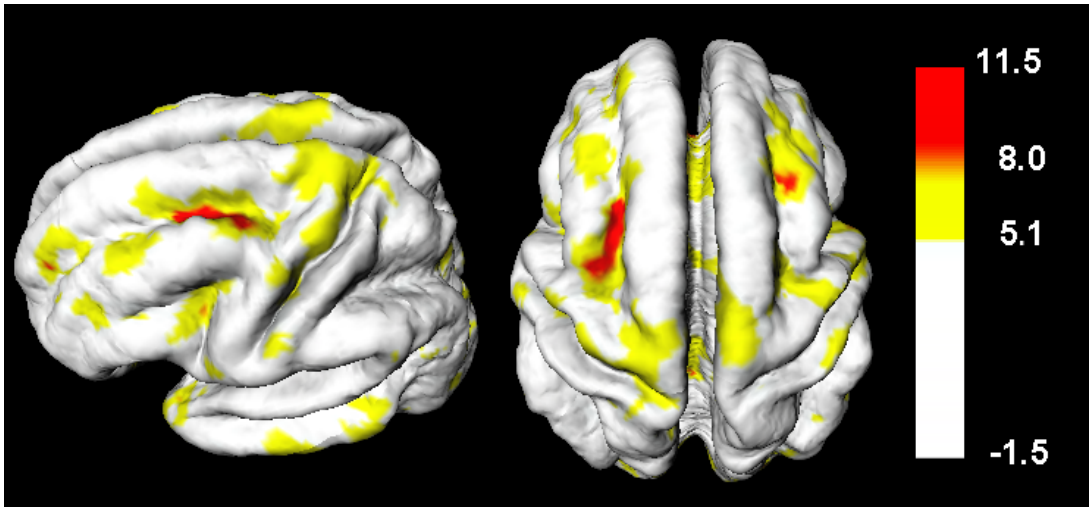

Figure 4.6: t map of the cortical thickness dilatation rate showing the statistically significant regions of cortical thickness increase (red & yellow). The red region exhibits extremely high t-values indicating that this is a region of extreme cortical thickness increase. There was no region of statistically significant cortical thinning showing that the cortical thickness tends to increase rather than decrease between ages 12 and 16.

where $\partial\Omega_{avg}$ is the outer cortical surface of the average atlas brain. There is a persistent global morphological pattern of cortical thickness increases by 2.6% per year and 11.3% over 4.6 year time span. We localized the region of statistically significant cortical thickness increase by thresholding the t-map of the cortical thickness dilatation rate by 5.1 (Figure 4.6). It is noted that there is *no* statistically significant cortical thinning detected on any region of the cortical surface. We conclude that, over all, gray matter gets thicker from ages 12 and 16. Further, there is an extremely localized region on the left hemisphere showing dominant cortical thickening as illustrated in Figure 4.6.

Therefore, the cortical surface area and gray-matter volume shrinks, while the cortical thickness tends to increase between ages 12 and 16 with a highly localized area of cortical thickening in the left hemisphere.

Subject	Age	$\|\partial\Omega_0^{out}\|$	$\|\partial\Omega_t^{out}\|$	$\|\partial\Omega_0^{in}\|$	$\|\partial\Omega_t^{in}\|$	$\|\Omega_0\|$	$\|\Omega_t\|$
1	14.05-19.13	2.8440	2.6279	2.2341	2.1342	7.4053	7.1212
2	12.73-17.05	2.9196	2.7315	2.1802	2.1237	7.3156	7.2197
3	10.02-16.41	3.0397	2.8584	2.3517	2.1665	7.6754	7.5205
4	13.94-18.66	3.0370	2.8911	2.3927	2.2740	7.8623	7.5949
5	13.62-18.40	2.9102	2.8473	2.1742	2.1951	7.5943	7.1588
6	16.41-20.99	2.7442	2.6063	2.1305	2.0991	7.3460	7.2084
7	17.07-21.75	2.9705	2.9194	2.2726	2.2517	7.6856	7.6156
8	14.87-21.20	2.9657	2.7553	2.4481	2.3758	7.7043	7.3001
9	8.97-14.86	3.2839	3.1537	2.3632	2.2858	8.1128	7.7807
10	9.85-15.93	3.0223	2.9745	2.4538	2.3637	7.8178	7.5884
11	16.41-21.47	2.8655	2.7851	2.2575	2.2151	7.3117	7.2331
12	8.28-13.18	3.1398	3.1078	2.2911	2.2150	7.8969	7.7978
13	8.45-13.09	2.9811	2.8123	2.2729	2.1097	7.4548	7.5112
14	11.82-16.08	2.9605	2.7772	2.3074	2.2053	7.7615	7.5062
15	9.18-13.29	3.0675	2.9695	2.2396	2.2056	7.8531	7.6688
16	12.72-17.01	2.7546	2.5915	2.0335	1.9352	7.5793	7.2319
17	8.15-12.66	3.0707	2.7841	2.2696	2.1558	7.8365	7.1807
18	9.48-13.93	3.0290	2.7762	2.1743	1.9861	7.6984	7.3084
19	7.72-12.67	3.2150	3.1199	2.4930	2.4298	7.7814	7.8242
20	11.67-15.64	3.0980	3.0895	2.5191	2.4574	7.4377	7.6239
21	11.75-15.85	2.9646	2.7631	2.2507	2.0337	7.5835	7.2833
22	7.02-11.28	3.1425	3.0747	2.3166	2.2015	7.9344	8.0303
23	11.61-16.04	3.0077	2.7930	2.2689	2.1558	7.7683	7.3123
24	11.88-15.93	3.3111	3.1970	2.4125	2.3790	8.2071	8.0173
25	10.25-14.77	3.2974	3.1423	2.4211	2.3037	7.8385	7.8556
26	17.84-21.21	2.8572	2.8156	2.1970	2.1790	7.4161	7.4805
27	9.15-12.71	2.9364	2.8679	2.3416	2.3394	7.4969	7.6005
28	8.41-10.61	3.1745	3.1437	2.3165	2.2476	7.6836	7.4952
Mean	11.55-16.14	3.0218	2.8920	2.2994	2.2152	7.6807	7.5025

Table 4.1: Total cortical surface area of 28 subjects ($\times 10^5 \text{mm}^2$). The last two columns are the total gray-matter volume ($\times 10^5 \text{mm}^3$).

Subject	Age	$\Theta_{total\ area}$	$\Lambda_{total\ area}$	Θ_{gray}	Λ_{gray}
1	14.05-19.13	-0.0760	-0.0150	-0.0384	-0.0076
2	12.73-17.05	-0.0644	-0.0149	-0.0131	-0.0030
3	10.02-16.41	-0.0596	-0.0093	-0.0202	-0.0032
4	13.94-18.66	-0.0480	-0.0102	-0.0340	-0.0072
5	13.62-18.40	-0.0216	-0.0045	-0.0573	-0.0120
6	16.41-20.99	-0.0503	-0.0110	-0.0187	-0.0041
7	17.07-21.75	-0.0172	-0.0037	-0.0091	-0.0019
8	14.87-21.20	-0.0709	-0.0112	-0.0525	-0.0083
9	8.97-14.86	-0.0396	-0.0067	-0.0409	-0.0070
10	9.85-15.93	-0.0158	-0.0026	-0.0293	-0.0048
11	16.41-21.47	-0.0280	-0.0055	-0.0108	-0.0021
12	8.28-13.18	-0.0102	-0.0021	-0.0125	-0.0026
13	8.45-13.09	-0.0566	-0.0122	0.0076	0.0016
14	11.82-16.08	-0.0619	-0.0145	-0.0329	-0.0077
15	9.18-13.29	-0.0320	-0.0078	-0.0235	-0.0057
16	12.72-17.01	-0.0592	-0.0138	-0.0458	-0.0107
17	8.15-12.66	-0.0933	-0.0207	-0.0837	-0.0186
18	9.48-13.93	-0.0835	-0.0188	-0.0507	-0.0114
19	7.72-12.67	-0.0296	-0.0060	0.0055	0.0011
20	11.67-15.64	-0.0028	-0.0007	0.0250	0.0063
21	11.75-15.85	-0.0680	-0.0166	-0.0396	-0.0097
22	7.02-11.28	-0.0216	-0.0051	0.0121	0.0028
23	11.61-16.04	-0.0714	-0.0161	-0.0587	-0.0133
24	11.88-15.93	-0.0345	-0.0085	-0.0231	-0.0057
25	10.25-14.77	-0.0470	-0.0104	0.0022	0.0005
26	17.84-21.21	-0.0145	-0.0043	0.0087	0.0026
27	9.15-12.71	-0.0233	-0.0066	0.0138	0.0039
28	8.41-10.61	-0.0097	-0.0044	-0.0245	-0.0111
Mean	11.55-16.14	-0.0432	-0.0094	-0.0230	-0.0050

Table 4.2: Total cortical surface area dilatation $\Theta_{total\ area}$ and its rate $\Lambda_{total\ area}$. The last two columns are the total gray-matter volume dilatation Θ_{gray} and its rate Λ_{gray}.

Chapter 5

Conclusions and Future Research

5.1 Conclusions

The deformation-based morphometry presented in this thesis can localize the regions where local volume growth or loss occurs over temporally varying brain morphology by measuring the rate of local volume changes. By using the displacement velocity instead of the displacement itself in detecting the anatomical changes, temporal variabilities in MR images for different age groups and different time intervals can be accounted for. Extending the concept of deformation-based morphometry in Euclidean space to non-Euclidean space, we have developed a complete statistical procedure for surface-based morphometry, which can be used in detecting gray-matter deformation and its outer and inner cortical boundary.

In brain imaging, smoothing can enhance signal-to-noise ratio, making functional and structural effects easier to detect. In 3D volumetric images of fMRI, PET and MRI, the standard smoothing technique is the Gaussian kernel smoothing. Therefore, it is natural to extend Gaussian kernel smoothing to 2D surface data. The most natural generalization of Gaussian kernel smoothing on a curved surface is via the diffusion equation based on the Laplace-Beltrami operator. We have developed two different approaches to diffusion smoothing: the parametric method and the finite element method. In some applications, the parametric method may be more suitable than the finite element method and vice versa. For the Riemannian metric $ds^2 = \sum_{i,j=1}^{n} g_{ij}\ du^i du^j$ on a Riemannian manifold M, the

Laplace-Beltrami operator Δ is defined by

$$\Delta F = \frac{1}{|g|^{1/2}} \sum_{i,j=1}^{n} \frac{\partial}{\partial u^i}\Big(|g|^{1/2} g^{ij} \frac{\partial F}{\partial u^j}\Big), \tag{5.1}$$

where $|g| = \det(g_{ij})$ and $g^{-1} = (g^{ij})$. Using the finite element method, we estimated (5.1) on the triangular mesh of the brain surfaces. Let $F(\mathbf{p}_i)$ be the data on the i-th node $\mathbf{p}_i$ in the triangular mesh. If $\mathbf{p}_1$,...,$\mathbf{p}_m$ are m-neighboring nodes around the central node $\mathbf{p}$, the Laplace-Beltrami operator is estimated as $\widehat{\Delta F}(\mathbf{p}) = \sum_{i=1}^{m} w_i\big(F(\mathbf{p}_i) - F(\mathbf{p})\big)$ with the weights $w_i = (\cot\theta_i + \cot\phi_i)/|T|$, where θ_i and ϕ_i are the two angles opposite to the edge connecting $\mathbf{p}_i$ and $\mathbf{p}$, and $|T|$ is the sum of the areas of the m-incident triangles at $\mathbf{p}$. Afterwards, the diffusion equation is solved via the finite difference scheme:

$$F(\mathbf{p}, t_{n+1}) = F(\mathbf{p}, t_n) + (t_{n+1} - t_n)\widehat{\Delta F}(\mathbf{p}, t_n)$$

with the initial condition $F(\mathbf{p}, t_0) = f(\mathbf{p})$ for each node $\mathbf{p}$ on the triangular mesh. After N-iterations, the diffusion smoothing is locally equivalent to Gaussian kernel smoothing with smoothing parameter $h = \sqrt{2}(t_N - t_0)^{1/2}$.

As an illustration, we have applied the methods to MR scans of 28 normal children and adolescents and detected regions of brain tissue growth or loss in both whole brain volume and on the cortical surface. Between age 12 and 16, it is found that the brain tissue growth occurs most rapidly on the somatosensory and motor cortex as well as in the isthmus and splenium of the corpus callosum. Applying the surface-based morphometry, we found no statistically significant cortical thinning process. Instead, there is a dominant global pattern of cortical thickness increase over time while the both inner and outer surface areas and the volume of the gray-matter decreases. Also we were able to localize the regions of surface area increase on the left hemisphere and surface area decrease on the right hemisphere.

Our unified statistical framework based on the deformation-based morphometry can be further used as a tool for future investigations of neurodevelopmental disorders where volumetric analysis would be relevant. It can also be applied to a general morphological studies, such as testing for structural shape differences between two different groups of subjects.

5.2 Future Research

5.2.1 Growth Curve Model

The linear growth model (2.2) is a special case of the full model (2.1). It would be interesting to develop nonlinear brain growth model. It requires combining both standard longitudinal analysis techniques [65] with random fields setting. Let $J(x,t)$ be the Jacobian at the position x and time t of the deformation

$$x \to x + U(x,t).$$

A proposed brain growth model is

$$\frac{\partial J}{\partial t}(x,t) = \sum_i \psi_i(x)\varphi_i(t) + \epsilon(x,t),$$

where $\epsilon(x,t)$ is a Gaussian random field and $\{\varphi_i\}$ are temporal basis functions. From the Karhunen-Loeve expansion in (2.16), the error term can be decomposed as

$$\epsilon(x,t) = \sum_i \epsilon_i(x)\phi_i(t),$$

where $\{\epsilon_i\}$ are independent Gaussian random fields and $\{\phi_i\}$ are orthonormal bases. Without loss of generality, we may let $\varphi_i = \phi_i$. Then we have

$$\frac{\partial J}{\partial t}(x,t) = \sum_i \psi_i(x)\phi_i(t) + \sum_i \epsilon_i(x)\phi_i(t).$$

where and $\epsilon(x)$ is a Gaussian random field. The problem is to estimate the unknown coefficient functions $\{\psi_i\}$ possibly by minimizing the mean squared error. Alternately, we can also model brain development using nonparmetric kernel smoothing techniques [65].

5.2.2 Membrane Spline Energy

In our thesis, morphological descriptors were based on length, area and volume changes, which directly measures the amount of brain tissue growth or loss. It is possible to develop more sophisticated morphological descriptor that measures completely different morphological properties in brain deformation. Consider two geometric objects Ω_1 and Ω_2 in $\mathbb{R}^N$ which have slight shape variations. We are interested in identifying the regions of maximum shape

differences between Ω_1 and Ω_2. Let $U(x) = (U_1(x), \cdots, U_N(x))^t$ be the displacement vector field from Ω_1 to Ω_2, whose components are assumed to follow a mean zero stationary Gaussian random fields. The deformation from Ω_1 to Ω_2 can be assumed to minimizes an associated *membrane spline energy* [45] given by

$$E(U) = \int_{\Omega_1} \sum_{i,j=1}^{N} \Big(\frac{\partial U_j}{\partial x_i}\Big)^2 \, dx. \tag{5.2}$$

For instance, the membrane spline energy in 2-dimension is

$$E(U) = \int_{\Omega_1} \Big[\Big(\frac{\partial U_1}{\partial x_1}\Big)^2 + \Big(\frac{\partial U_2}{\partial x_1}\Big)^2 + \Big(\frac{\partial U_1}{\partial x_2}\Big)^2 + \Big(\frac{\partial U_2}{\partial x_2}\Big)^2\Big] \, dx_1 \, dx_2$$

The functional inside the energy integral in (5.2) is the squared *Frobenius norm* of the displacement gradient matrix ∇U in (2.4). Let us denote $\|\nabla U\|_F$ to be the Frobenius norm of ∇U, i.e.

$$\|\nabla U\|_F = \Big[\sum_{i,j=1}^{N} \Big(\frac{\partial U_j}{\partial x_i}\Big)^2\Big]^{1/2}.$$

The Frobenius norm will measure the amount of local membrane spline energy associated with the deformation $x \to x + U$ in the neighborhood of x and it would be very interesting to compare this local energy functional to local volume change and local displacement change statistics. Other spline energy functionals such as the *thin-plate splines* [12, 49, 115] might be also used as morphological descriptors.

5.2.3 Vorticity Tensor Fields

We introduced the concept of *vorticity tensor* in (2.10) which measure the amount of rotation at a given point but never studied its statistical properties. Assume an object Ω_1 to be at time 0 and after a unit time, the object Ω_1 goes through the viscous fluid deformation to Ω_2. The vorticity vector is defined as the curl of the velocity,

$$\omega(x) = \frac{1}{2} \nabla \times U(x),$$

where $\nabla = (\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_N})'$ and it completely determines infinitesimal rotation of the deformation. The obvious testing procedure for detecting any local rotational change is to use the Hotelling's T^2 statistics. Because the components of $\omega(x)$ are correlated, the random

field constructed from $\omega(x)$ is not the Hotelling's T^2 field. So it would be of interest to be able to compute the excursion probability based on this Hotelling's T^2 like random field.

The vorticity tensor ω_{ij} is given by

$$\omega_{ij} = \frac{1}{2}\Big(\frac{\partial U_i}{\partial x_j} - \frac{\partial U_j}{\partial x_i}\Big).$$

Let ϵ_{ijk} be the *Levi-Civita tensor* [94, 75], then the i-th element of ω can be given in terms of the vorticity tensor

$$\omega_i = -\epsilon_{ijk}\omega_{jk}.$$

The angular speed is defined as

$$\|\omega\| = \big(\sum_{i=1}^{N} \omega_i^2\big)^{1/2}.$$

The angular speed is a useful scalar morphological descriptor which measures the amount of rotation per unit time in deformation. Finding the exact statistical distribution and its p-value based on the maximum of its field seems somewhat complicated.

5.2.4 Generalized Variance Field

In (2.4), we defined the 3-dimensional displacement gradient matrix. We introduce a new morphological descriptor based on the determinant of the matrix. Suppose components of the displacement vector $U(x) \in \mathbb{R}^N$ are identically and independently distributed as a mean zero stationary Gaussian field with the covariance function $R(x, y) = f(x - y)$. Let

$$\nabla U_i = \Big(\frac{\partial U_i}{\partial x_1}, \cdots, \frac{\partial U_i}{\partial x_N}\Big)^t$$

and ∇U be the N-dimensional *displacement gradient matrix* defined by

$$\nabla U = \big(\nabla U_1, \cdots, \nabla U_N\big)^t = \begin{pmatrix} \frac{\partial U_1}{\partial x_1} & \cdots & \frac{\partial U_1}{\partial x_N} \\ & \ddots & \\ \frac{\partial U_N}{\partial x_1} & \cdots & \frac{\partial U_N}{\partial x_N} \end{pmatrix}.$$

The *generalized variance field* is defined as the determinant of $W = (\nabla U)^t \nabla U$. Note that

$$W(x) = \sum_{i=1}^{N} \nabla U_i (\nabla U_i)^t.$$

It would be very useful to approximate the excursion probability of the generalized variance field. Unfortunately, it is not easy to compute the expected Euler characteristic of the excursion set of W.

The covariance function of the field $W^{1/2}$ can be easily computed. By expanding the determinant $W^{1/2} = \det(\nabla U)$,

$$W^{1/2} = \sum_{\sigma \in S_N} \mathrm{sgn}(\sigma) \frac{\partial U_1}{\partial x_{\sigma(1)}} \cdots \frac{\partial U_N}{\partial x_{\sigma(N)}},$$

where S_N is a symmetric group of order N and $\mathrm{sgn}(\sigma)$ is the sign function of the order of permutation σ [54]. Since $\frac{\partial U_1}{\partial x_{\sigma(1)}}, \ldots, \frac{\partial U_N}{\partial x_{\sigma(N)}}$ are independent mean zero Gaussian fields, $\mathbb{E}(W^{1/2}) = 0$. The covariance function R^* of $W^{1/2}$ is

$$\begin{aligned}
R^*(x,y) &= \mathbb{E}[W^{1/2}(x) W^{1/2}(y)] \\
&= \mathbb{E}\left[\sum_{\sigma,\tau \in S_N} \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau) \frac{\partial U_1(x)}{\partial x_{\sigma(1)}} \frac{\partial U_1(y)}{\partial y_{\tau(1)}} \cdots \frac{\partial U_N(x)}{\partial x_{\sigma(N)}} \frac{\partial U_N(y)}{\partial y_{\tau(N)}} \right] \\
&= \sum_{\sigma,\tau \in S_N} \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau) \mathbb{E}\left[\frac{\partial U_1(x)}{\partial x_{\sigma(1)}} \frac{\partial U_1(y)}{\partial y_{\tau(1)}} \right] \cdots \mathbb{E}\left[\frac{\partial U_N(x)}{\partial x_{\sigma(N)}} \frac{\partial U_N(y)}{\partial y_{\tau(N)}} \right] \\
&= \sum_{\sigma,\tau \in S_N} \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau) R_{\sigma(1)\tau(1)}(x,y) \cdots R_{\sigma(N)\tau(N)}(x,y),
\end{aligned}$$

where $R_{ij}(x,y) = \frac{\partial^2}{\partial x_i \partial y_i} R(x,y)$. There exists a permutation $\rho \in S_N$ such that $\tau = \rho\sigma$ and $\mathrm{sgn}(\rho) = \mathrm{sgn}(\tau)\mathrm{sgn}(\sigma)$. Then by summing up over the index ρ,

$$\begin{aligned}
R^*(x,y) &= \sum_{\sigma \in S_N} \sum_{\rho \in S_N} \mathrm{sgn}(\rho) R_{\sigma(1)\rho\sigma(1)}(x,y) \cdots R_{\sigma(N)\rho\sigma(N)}(x,y) \\
&= \sum_{\sigma \in S_N} \det\left(\frac{\partial^2}{\partial x \partial y^t} R(x,y) \right) \\
&= N! \det\left(\frac{\partial^2}{\partial x \partial y^t} R(x,y) \right) \\
&= (-1)^N \det\Big(\mathbf{H} f(x-y) \Big),
\end{aligned}$$

where $\mathbf{H}f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)$ is the Hessian matrix of f.

Note

The Matlab program used in computing the corrected thresholds of the T random field and the Hotelling's T^2 random field can be found at

`http://www.math.mcgill.ca/keith/BICstat`

The Matlab program used in diffusion smoothing can be found at

`http://www.math.mcgill.ca/chung/diffusion`

Bibliography

[1] R.J. Adler. *The Geometry of Random Fields*. John Wiley & Sons, 1981.

[2] R.J. Adler. *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes*. IMS, Hayward, CA, 1990.

[3] D. Aldous. *Probability Approximations via the Poisson Clumping Heuristic*. Springer-Verlag, New Yor, 1989.

[4] A. Andrade, F. Kherif, J. Mangin, K.J. Worsley, A. Paradis, O. Simon, S. Dehaene, D. Le Bihan, and J-B Poline. Detection of FMRI activation using cortical surface mapping. *Human Brain Mapping*, 12:79–93, 2001.

[5] E. Angelopoulou. Gaussian curvatures from photometric scatter plots. *Workshop on Photometric Modeling for Computer Vision and Graphics*, pages 12–19, 1999.

[6] S. Angenent, S. Hacker, A. Tannenbaum, and R. Kikinis. On the Laplace-Beltrami operator and brain surface flattening. *IEEE Transactions on Medical Imaging*, 18:700–711, 1999.

[7] G.B. Arfken. *Mathematical Methods for Physicists*. Academic Press, 5th. edition, 2000.

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

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

[10] M. Bakircioglu, S. Joshi, and M.I. Miller. Landmark matching on the sphere via large deformation diffeomorphisms. *Proceedings of SPIE Medical Imaging 1999: Image Processing*.

[11] I.V. Basawa and B.L.S.P. Rao. *Statistical Inference for Stochastic Processes*. Academic Press, London, 1980.

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

[13] F.L. Bookstein. *Morphometric Tools for Landmark Data:Geometry and Biology*. Cambridge University Press, Cambridge, 1991.

[14] W.M. Boothby. *An Introduction to Differential Manifolds and Riemannian Geometry*. Academic Press, London, second edition edition, 1986.

[15] A.W. Bowman and A. Azzalini. *Applied Smoothing Techniques for Data Analysis*. Oxford University Press Inc., New York, 1997.

[16] J. Cao. The size of the connected components of excursion sets of χ^2, t and f fields. *Advances in Applied Probability*, 1999.

[17] J. Cao and K.J. Worsley. The detection of local shape changes via the geometry of hotelling's t^2 fields. *Annals of Statistics*, 27:925–942, 1999.

[18] M.P. Do Carmo. Riemannian geometry. *Prentice-Hall, Inc.*, 1992.

[19] D.S. Chandrasekharaiah and L. Debnath. *Continuum Mechanics*. Academic Press, San Diego, 1994.

[20] G.E. Christensen, S.C. Joshi, and M.I Miller. Volumetric transformation of brain anatomy. *IEEE Transactions on Medical Imaging*, 16:864–877, 1997.

[21] G.E. Christensen, R.D. Rabbitt, and M.I Miller. A deformable neuroanatomy textbook based on viscous fluid mechanics. *Proc. of the 27th Annual Conference on Information Science and Systems*, pages 211–216, 1993.

[22] M.K. Chung, K.J. Worsley, C. Cherif, T. Paus, D.L. Collins, J. Giedd, J.L. Rapoport, and A.C. Evans. Statistical analysis of local volume change, with an application to brain growth. *NeuroImage*, 11:S611, 2000.

[23] M.K. Chung, K.J. Worsley, T. Paus, D.L. Cherif, C. Collins, J. Giedd, J.L. Rapoport, and A.C. Evans. A unified statistical approach to deformation-based morphometry. *NeuroImage*, 14, 2001.

[24] M.K. Chung, K.J. Worsley, S. Robbins, T. Paus, J. Taylor, J.N. Giedd, J.L. Rapoport, and A.C. Evans. Statistical analysis of cortical surface area change, with an application to brain growth. *Proc. of the 7th Annual Meeting of the Organization for Human Brain Mapping*, 2001.

[25] D.L. Collins and A.C Evans. Animal:automatic nonlinear image matching and anatomical labeling. *Brain Warping*, pages 133–142, 1999.

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

[27] D.L. Collins, T. Paus, A. Zijdenbos, K.J. Worsley, J. Blumenthal, J.N. Giedd, and A.C. Rapoport, J.L.and Evans. Age related changes in the shape of temporal and frontal lobes: an mri study of children and adolescents. *Soc. Neuroscience Abstract*, 24:304, 1998.

[28] R. Courant. Variational methods for the solution of problems of equilibrium and vibrations. *Bulletin of American Mathematical Society*, 49, 1943.

[29] R. Courant and D. Hilbert. *Methods of Mathematical Physics: Vol. II.* English edition, 1953.

[30] E. Courchesne, H.J. Chisum, J. Townsend, A. Cowles, J. Covington, B. Egaas, M. Harwood, S. Hinds, and G.A. Press. Normal brain development and aging : Quantitative analysis at in vivo mr imaging in healthy volunteers. *Radiology*, 216:672–682, 2000.

[31] A.M Dale and B. Fischl. Cortical surface-based analysis I. segmentation and surface reconstruction. *NeuroImage*, 9:179–194, 1999.

[32] C. Davatzikos. Brain morphometrics using geometry-based shape transformations. *Proc. of Workshop on Biomed. Image Registration, Slovenia*, 1999.

[33] C. Davatzikos and R.N. Bryan. Using a deformable surface model to obtain a shape representation of the cortex. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2122–2127, 1995.

[34] C. Davatzikos, M. Vaillant, S.M. Resnick, J.L. Prince, S. Letovsky, and N. Bryan. A computerized approach for morphological analysis of the corpus callosum. *Journal of Computer Assisted Tomography*, 20:88–97, 1996.

[35] A.S. Dekaban. Tables of cranial and orbital measurements cranial volume and derived indexes in males and females from 7 days to 20 years of age. *Ann. Neurology*, 2:485–491, 1977.

[36] A.S. Dekaban and D. Shadowsky. Changes in brain weights during the span of human life: relation of brain weights to body heights and body weights. *Ann. Neurology*, 4:345–356, 1978.

[37] M. Desbrun, M. Meyer, P. Schroder, and A.H. Barr. Implicit fairing of irregular meshes using diffusion and curvature flow. *Computer Graphics (SIGGRAPH 99 Proceedings)*, pages 317–324, 1999.

[38] E.R. Dougherty. *Random Processes for Image and Signal Processing.* IEEE Press, 1999.

[39] D.A. Drew. *Theory of multicomponent fluids.* Springer-Verlag, New York, 1991.

[40] I.L. Dryden and K.V. Mardia. *Statistical Shape Analysis.* John Wiley & Sons, 1998.

[41] Y.V. Egorov and M.A. Shubin. *Partial Differential Equations I: Foundations of the Classical Theory.* Springer-Verlag, 1991.

[42] G. Gallo, A. Zingale, and R. Zingale. Detection of mri brain contour using nonlinear anisotropic diffusion filter, bridging disciplines for biomedicine. *18th Annual International Conference of the IEEE*, 3:1062–1064, 1997.

[43] C. Gaser, H.-P. Volz, S. Kiebel, and H. Riehemann, S.and Sauer. Detecting structural changes in whole brain based on nonlinear deformations - application to schizophrenia research. *NeuroImage*, 10:107–113, 1999.

[44] J.C. Gee and R.K. Bajcsy. Elastic matching: Continuum mechanical and probabilistic analysis. *Brain Warping*, pages 183–198, 1999.

[45] J.C. Gee, M. Reivich, and R. Bajcsy. Elastically deforming an atlas to match anatomical brain images. *Journal of Computer Assisted Tomography*, 17:225–236, 1993.

[46] J.N. Giedd, J. Blumenthal, N.O. Jeffries, J.C. Rajapakse, A.C. Vaituzis, H. Liu, Y.C. Berry., M. Tobin, J. Nelson, and F.X. Castellanos. Development of the human corpus callosum during childhood and adolescence: a longitudinal mri study. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 23:571–588, 1999.

[47] J.N. Giedd, J.W. Snell, N. Lange, J.C. Rajapakse, D. Kaysen, A.C. Vaituzis, Y.C. Vauss, S.D. Hamburger, P.L. Kozuch, and J.L. Rapoport. Quantitative magnetic resonance imaging of human brain development: Ages 4-18. *Cerebral Cortex*, 6:551–160, 1996.

[48] R. Goebel and W. Singer. Cortical surface-based statistical analysis of functional activation. *NeuroImage*, 9S:64, 1999.

[49] P.G. Green and B.W. Silverman. *Non-parametric Regression and Generalized Linear Models: A Roughness Penalty Approach.* Chapman and Hall, London, 1994.

[50] L.D. Griffin. The intrinsic geometry of the cerebral cortex. *Journal of Theoretical Biology*, 166:261–273, 1994.

[51] A. Grigoryan. Analytic and geometric background of recurrence and non-explosion of the brownian motion on riemannian manifolds. *Bulletin of the American Mathematical Society*, 36:135–249, 1999.

[52] D. Guéziec. *Spline Curves and Surfaces for Data Modeling, Advances in Morphometrics.* Plenum Press, New Work, 1996.

[53] M.E. Gurtin and G.B. McFadden. *On the Evolutation of Phase Boundaries.* Springer-Verlag, New York, 1991.

[54] D.A. Harville. *Matrix Algebra from a Statistician's Perspective.* Springer-Verlag, New York, 1997.

[55] C. Henery and T. Mayhew. The cerebrum and cerebellum of the fixed human brain: efficient and unbiased estimates of volumes and cortical surface areas. *J. Anat.*, 167:167–180, 1989.

[56] G. Huiskamp. Difference formulas for the surface laplacian on a triangulated surface. *Journal of Computational Physics*, 95:477–496, 1991.

[57] T.L. Jernigan, D.A. Trauner, J.R. Hesselink, and P.A. Tallal. Maturation of human cerebrum observed in vivo during adolescence. *Brain*, 114:2037–2049, 1991.

[58] S.E. Jones, B. R. Buchbinder, and I. Aharon. Three-dimensional mapping of cortical thickness using laplace's equation. *Human Brain Mapping*, 11:12–32, 2000.

[59] S.C. Joshi. *Large Deformation Diffeomorphisms and Gaussian Random Fields for Statistical Characterization of Brain Sub-Manifolds.* Ph.D. Thesis, Washington University, St. Louis., 1998.

[60] S.C. Joshi, J. Wang, M.I. Miller, D.C. Van Essen, and U. Grenander. On the differential geometry of the cortical surface. *Vision Geometry IV, Vol. 2573, Proceedings of the SPIE's 1995 International Symposium on Optical Science, Engineering and Instrumentation*, pages 304–311, 1995.

[61] N. Kabani, G. Le Goualher, D. MacDonald, and A. C. Evans. Measurement of cortical thickness using an automated 3-d algorithm: A validation study. *NeuroImage*, 13:375–380, 2000.

[62] D.G. Kendall. A survey of the statistical theory of shape. *Statistical Science*, 4:87–120, 1989.

[63] S. Kovačič and R. Bajcsy. Multiscale/multiresolution representations. *Brain Warping*, pages 45–65, 1999.

[64] E. Kreyszig. *Differential Geometry.* University of Toronto Press, Toronto, 1959.

[65] A.M. Kshirsagar and W.B. Smith. Growth curves. Marcel Dekker, Inc., 1995.

[66] S. Kwapien and W.A. Woyczynski. *Random Series and Stochastic Integrals: Single and Multiple.* Probability and Its Applications. Birkhauser, 1992.

[67] L.D. Landau and E.M. Lifshitz. *Fluid Mechanics*. Pagamon Press, 2nd. edition edition, 1989.

[68] C.L. Lawson and R.J. Hanson. *Solving Least Sqaures Problems*. Prentice-Hall, 1974.

[69] M.R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag, New York, 1982.

[70] A. Lester and R.A. Simon. A survey of hierarchical non-linear medical image registration. *Pattern recognition*, 32:129–149, 1999.

[71] T. Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publisher, 1994.

[72] W.E. Lorensen and H.E. Cline. Marching cubes: A hight resolution 3d surface construction algorithm. *Computer Graphics*, 21:21, 1987.

[73] J.D. MacDonald. *A Method for Identifying Geometrically Simple Surfaces from Three Dimensional Images*. PhD thesis, McGill University, Montreal, Canada, 1997.

[74] J.D. MacDonald, N. Kabani, D. Avis, and A.C. Evans. Automated 3-d extraction of inner and outer surfaces of cerebral cortex from mri. *NeuroImage*, 12:340–356, 2000.

[75] J.E. Marsden and T.J.R. Hughes. *Mathematical Foundations of Elasticity*. Dover Publications, Inc., 1983.

[76] R.J.Y. McLeod and M.L. Baart. *Geometry and Interpolation of Curves and Surfaces*. Cambridge University Press, 1998.

[77] J. Oden. *Finite Elements: Mathematical Aspects*. Prentice-Hall, New Jersey, 1983.

[78] Y. Ohtake, A.G. Belyaev, and I.A. Bogaevski. Polyedral surface smoothing with simultaneous mesh regularization. *Proceedings of Geometric Modeling and Processing. Theory and Applications*, pages 229–237, 2000.

[79] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

[80] A. Pfefferbaum, D.H. Mathalon, E.V. Sullivan, J.M. Rawles, and K.O. Zipursky, R.B.and Lim. A quantitative magnetic resonance imaging study of changes in brain morphology from infancey to late adulthood. *Archives of Neurology*, 51:874–887, 1994.

[81] J-B Poline and B.M. Mazoyer. Analysis of individual brain activation maps using hierarchical description and multiscale detection. *IEEE Transactions on Medical Imaging*, 13:702–710, 1994.

[82] J-B Poline and B.M. Mazoyer. Enhanced detection in activation maps using a multifiltering approach. *Journal of Cerebral Blood Flow and Metabolism*, 14:690–699, 1994.

[83] J-B Poline, K.J. Worsley, Frackowiak R.S.J. Holmes, A.P., and K.J. Friston. Estimating smoothness in statistical parametric maps: Variability of p values. *Journal of Computer Assisted Tomography*, 19:788–796, 1995.

[84] J. Pujol, P. Vendrell, C. Junque, J.M. Martivilalta, and A. Capdevila. When does human brain development end? evidence of corpus callosum growth up to adulthood. *Annals of Neurology*, 34:71–75, 1993.

[85] J.C. Rajapakse, J.N. Giedd, C. DeCarli, J.W. Snell, A. McLaughlin, Y.C. Vauss, A.L. Krain, S. Hamburger, and J.L. Rapoport. A technique for single-channel mr brain tissue segmentation: Application to a pediatric sample. *Magnetic Resonance Imaging*, 14:1053–1065, 1996.

[86] J.O. Ramsay. Differential equation models for statistical functions. *The Canadian Journal of Statistics*, 28:225–240, 2000.

[87] J.O. Ramsay and B.W. Silverman. *Functional Data Analysis*. Springer-Verlag, New York, 1997.

[88] T. Ramsay. *A Bivariate Finite Element Smoothing Spline Applied to Image Registration*. PhD thesis, Queen's University, Kingston, Canada, 1999.

[89] A.L. Riess, M.T. Abrams, H.S. Singer, J.L. Ross, and M.B. Denckla. Brain development, gender and iq in children: A volumetric imaging study. *Brain*, 119:1763–1774, 1996.

[90] B.D. Ripley. *Spatial Statistics*. John Wiley & Sons, Inc, 1981.

[91] A. Rosenfeld and A.C. Kak. *Digital Picture Processing*, volume Volume 2. Academic Press, Orlando, 1982.

[92] M.N.O. Sadiku. A simple introduction to finite element analysis of electromagnetic problems. *IEEE Transactions on Education*, 32:85–93, 1989.

[93] M.N.O. Sadiku. *Numerical Techniques in Electromagnetics*. CRC Press, 1992.

[94] P.G. Saffman. *Vortex Dynamics*. Cambridge University Press, New York, 1992.

[95] F.E. Satterthwaite. An approximate distribution of estimates of variance components. *Biometrics*, 2:110–114, 1946.

[96] D. W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons, Inc, 1992.

[97] J.A. Sethian. *Level Set Methods and Fast Marching Methods*. Cambridge University Press, 1996.

[98] D.O. Siegmund and K.J. Worsley. Testing for a signal with unknown location and scale in a stationary gaussian random field. *Annals of Statistics*, 23:608–639, 1995.

[99] P.P. Silvester and R.L. Ferrari. *Finite Elements for Electrical Engineers*. Cambridge: Cambridge Univ. Press, 1983.

[100] J.S. Simonoff. *Smoothing Methods in Statistics*. Springer, New York, 1996.

[101] C.G. Small. *The Statistical Theory of Shape*. Springer, New York, 1996.

[102] J.C. Strikwerda. *Finite Difference Schemes and Partial Differential Equations*. Wadsworth, Inc., Belmont, 1989.

[103] G. Subsol. Crest lines for curve-based warping. *Brain Warping*, pages 241–262, 1999.

[104] J. Talairach and P. Tournoux. Co-planar stereotactic atlas of the human brain: 3-dimensional proportional system. an approach to cerebral imaging. *Thieme, Stuttgart*, 1988.

[105] G. Taubin. Curve and surface smoothing without shrinkage. *Proc. Fifth Int. Conf. Computer Vision*, pages 852–857, 1995.

[106] J.E. Taylor and R.J. Adler. Euler characteristics for gaussian fields on manifolds. *Annals of Probability*, 2001. submitted.

[107] J.-P. Thirion and G. Calmon. Deformation analysis to detect quantify active lesions in 3d medicl image sequences. *IEEE Transactions on Medical Imaging*, 18:429–441, 1999.

[108] T.Y. Thomas. *Concepts from Tensor Analysis and Differential Geometry.* Academic Press, 1961.

[109] D.W. Thompson. *On Growth and Form.* Cambridge University Press, New York, 1961.

[110] P.M. Thompson, J.N. Giedd, R.P. Woods, D. MacDonald, A.C. Evans, and A.W. Toga. Growth patterns in the developing human brain detected using continuum-mechanical tensor mapping. *Nature*, 404:190–193, 2000.

[111] P.M. Thompson, D. MacDonald, M.S. Mega, C.J. Holmes, A.C. Evans, and A.W. Toga. Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces. *Journal of Computer Assisted Tomography*, 21:567–581, 1997.

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

[113] P.M. Thompson and A.W. Toga. Anatomically driven strategies for high-dimensional brain image warping and pathology detection. *Brain Warping*, pages 311–336, 1999.

[114] A.W. Toga. *Brain Warping.* Academic Press, 1999.

[115] G. Wahba. *Spline Models for Observational Data.* SIAM, Philadelphia, 1990.

[116] S. Warfield, A. Robatino, J. Dengler, F. Jolesz, and R. Kikinis. Brain warping. 1999.

[117] A. Witkin. Scale-space filtering. *Int. Joint Conf. Artificial Intelligence*, pages 1019–1021, 1983.

[118] K.J. Worlsey, J-B. Poline, A.C. Vandal, and K.J. Friston. Test for distributed, non-focal brain activations. *NeuroImage*, 2:173–181, 1995.

[119] K. J. Worsley. Testing for signals with unknown location and scale in a χ^2 random field, with an application to fmri. *Advances in Applied Probability*, 1998. in press.

[120] K. J. Worsley, M. Andermann, T. Koulis, D. MacDonald, and A. C. Evans. Detecting changes in non-isotropic images. *Human Brain Mapping*, 8:98–101, 1999.

[121] K.J. Worsley. Local maxima and the expected euler characteristic of excursion sets of χ^2, f and t fields. *Advances in Applied Probability*, 26:13–42, 1994.

[122] K.J. Worsley. *An unbiased estimator for the roughness of a multivariate Gaussian random field: Technical Report.* Department of Mathematics and Statistics, McGill University, 1996.

[123] K.J. Worsley, S. Marrett, P. Neelin, and A.C. Evans. A three-dimensional statistical analysis for cbf activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, 12:900–918, 1992.

[124] K.J. Worsley, S. Marrett, P. Neelin, and A.C. Evans. Searching scale space for activation in pet images. *Human Brain Mapping*, 4:74–90, 1996.

[125] K.J. Worsley, S. Marrett, P. Neelin, A.C. Vandal, K.J. Friston, and A.C. Evans. A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, 4:58–73, 1996.

[126] A.M. Yaglom. Correlation theory of stationary and related random functions vol. i: Basic results. 1987.

[127] E. Zauderer. *Partial Differential Equations of Applied Mathematics.* John Wiley & Sons, 2nd. edition, 1989.

[128] A.P. Zijdenbos, A. Jimenez, and A.C Evans. Pipelines: Large scale automatic analysis of 3D brain data sets. *NeuroImage*, 7S:783, 1998.