9

Statistical Analysis on Brain Surfaces

Moo K. Chung and Seth D. Pollak

University of Wisconsin-Madison

Jamie L. Hanson

Duke University

CONTENTS

In this chapter, we review widely used statistical analysis frameworks for data defined along cortical and subcortical surfaces that have been developed in last two decades. The cerebral cortex has the topology of a 2D highly convoluted sheet. For data obtained along curved non-Euclidean surfaces, traditional statistical analysis and smoothing techniques based on the Euclidean metric structure are inefficient. To increase the signal-to-noise ratio (SNR) and to boost the sensitivity of the analysis, it is necessary to smooth out noisy surface data. However, this requires smoothing data on curved cortical manifolds and assigning smoothing weights based on the geodesic distance along the surface. Thus, many cortical surface data analysis frameworks are differential geometric in nature [2]. The smoothed surface data is then treated as smooth random fields and statistical inferences can be performed within Keith Worsley's random field theory [3, 4].

The methods described in this chapter are illustrated with the hippocampus surface data set published in [5]. Using this case study, we will determine if there is an effect of

family income on the growth of hippocampus in children in detail. There are a total of 124 children and 82 of them have repeat magnetic resonance images (MRIs) two years later.

9.1   Introduction

The cerebral cortex has the topology of a 2D convoluted sheet. Most of the features that distinguish these cortical regions can only be measured relative to the local orientation of the cortical surface [6]. As the brain develops over time, cortical surface area expands and its curvature changes [2]. It is equally likely that such age-related changes with respect to the cortical surface are not uniform [7, 8]. By measuring how geometric features such as the cortical thickness, curvature and local surface area change over time, statistically significant brain tissue growth or loss in the cortex can be detected locally at the vertex level.

The first obstacle in performing surface-based data analysis is a need for extracting cortical surfaces from MRI volumes. This requires correcting MRI field inhomogeneity artifacts. The most widely used technique is the nonparametric nonuniform intensity normalization method (N3) developed at the Montreal neurological institute (MNI), which eliminates the dependence of the field estimate on anatomy [9]. The next step is the tissue classification into three types: gray matter, white matter and cerebrospinal fluid (CSF). This is critical for identifying the tissue boundaries where surface measurements are obtained. An artificial neural network classifier [10, 11] or Gaussian mixture models [12] can be used to segment the tissue types automatically. The Statistical Parametric Mapping (SPM) package[1] uses a Gaussian mixture with a prior tissue density map.

After the segmentation, the tissue boundaries are extracted as triangular meshes. In order to triangulate the boundaries, the marching cubes algorithm [13], level set method [14], the deformable surfaces method [15] or anatomic segmentation using proximities (ASP) method [16] can be used. Brain substructures such as the brain stem and the cerebellum are usually automatically removed in the process. The resulting triangular mesh is expected to be topologically equivalent to a sphere. For example, the triangular mesh resulted from the ASP method consists of 40,962 vertices and 81,920 triangles with the average internodal distance of 3 mm. Figure 9.1 shows a representative cortical mesh obtained from ASP. Surface measurements such as cortical thickness can be automatically obtained at each mesh vertex. Subcortical brain surfaces such as amygdala and hippocampus are extracted similarly, but often done in a semi-automatic fashion with the marching cubes algorithm on manual edited subcortical volumes. In the hippocampus case study, the left and right hippocampi were manually segmented in the template using the protocol outlined in [17].

Comparing measurements defined across different cortical surfaces is not trivial due to the fact that no two cortical surfaces are identically shaped. In comparing measurements across different 3D whole brain images, 3D volume-based image registration such as Advanced Normalization Tools (ANTS) [1] is needed. However, 3D image registration techniques tend to misalign sulcal and gyral folding patterns of the cortex. Hence, 2D surface-based registration is needed in order to match measurements across different cortical surfaces. Various surface registration methods have been proposed [18, 2, 19, 20, 21, 22]. Most methods solve a complicated optimization problem of minimizing the measure of discrepancy between two surfaces. A much simpler spherical harmonic representation technique provide a simple way of approximately matching surfaces without time-consuming numerical optimization [2].

---

[1]The SPM package is available at www.fil.ion.ucl.ac.uk/spm.
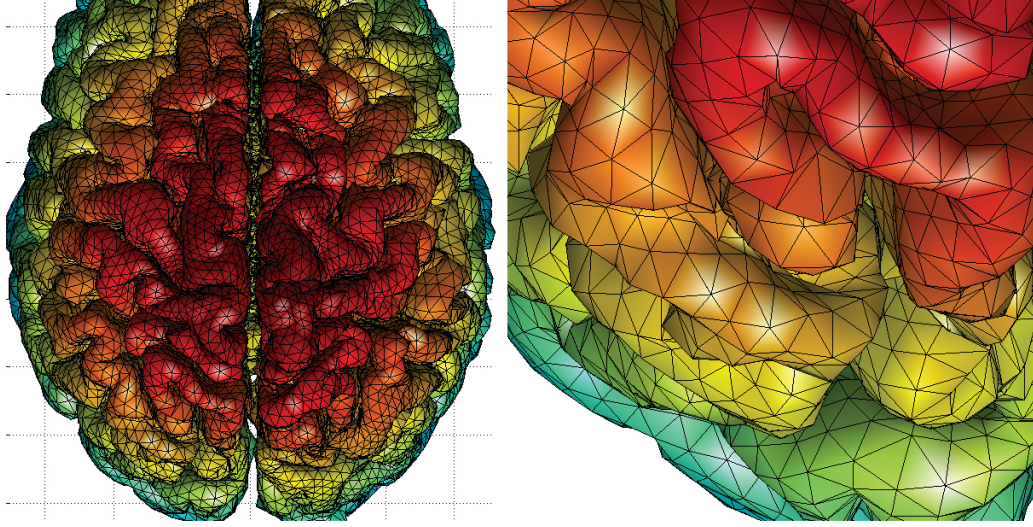
FIGURE 9.1
Left: The outer cortical brain surface mesh consisting of 81,920 triangles. Measurements are defined at mesh vertices. Right: The part of the mesh is enlarged to show the convoluted nature of the surface.

Surface registration and the subsequent surface-based analysis usually require parameterizing surfaces. It is natural to assume the surface mesh to be a discrete approximation to the underlying cortical surface, which can be treated as a smooth 2D Riemannian manifold. Cortical surface parameterization has been done previously in [19, 23, 2]. The surface parameterization also provides surface shape features such as the Gaussian and mean curvatures, which measure anatomical variations associated with the deformation of the cortical surface during, for instance, development and aging [6, 24, 23, 2].

9.2 Surface Parameterization

In order to perform statistical analysis on a surface, parameterization of the surface is often required [2]. Brain surfaces are often mapped onto a plane or a sphere. Then surface measurements defined on mesh vertices are also mapped onto the new domain and analyzed. However, almost all surface parameterizations suffer metric distortions, which in turn influence the spatial covariance structure so it is not necessarily the best approach.

We model the cortical surface $\mathcal{M}$ as a smooth 2D Riemannian manifold parameterized by two parameters u^1 and u^2 such that any point $\mathbf{x} \in \mathcal{M}$ can be represented as

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

for some parameter space $\mathbf{u} = (u^1, u^2) \in D \subset \mathbb{R}^2$ [25, 26, 7, 27]. The aim of the parameterization is estimating the coordinate functions x_1, x_2, x_3 as smoothly as possible.

Both global and local parameterizations are available. A global parameterization, such as tensor B-splines and spherical harmonic representation, are computationally expensive compared to a local surface parameterization. A local surface parameterization in the neigh-

borhood of point $\mathbf{x} = (x_1, x_2, x_3)$ can be obtained via the projection of the local surface patch onto the tangent plane $T_{\mathbf{x}}(\mathcal{M})$ [2, 23].

9.2.1 Local Parameterization by Quadratic Polynomial

A local parameterization is usually done by fitting a quadratic polynomial of the form

$$\mathbf{X}(u^1, u^2) = \beta_1 u^1 + \beta_2 u^2 + \beta_3 (u^1)^2 + \beta_4 u^1 u^2 + \beta_5 (u^2)^2 \tag{9.1}$$

in $(u^1, u^2) \in D \subset \mathbb{R}^2$. The data can be centered so there is no constant term in the quadratic form (9.1) [2]. The coefficients β_i are usually estimated by the least squares method [23, 28, 2].

In estimating various differential geometric measures such as the Laplace-Beltrami operator or curvatures, it is not necessary to find global surface parameterization of $\mathcal{M}$. Local surface parameterization such as the quadratic polynomial fit is sufficient to obtain such geometric quantities [2]. The drawback of the polynomial parameterization is that there is a tendency to weave the outermost mesh vertices to find vertices in the center. Therefore this is not advisable to directly fit (9.1) when one of the coordinate values rapidly changes.

9.2.2 Surface Flattening

Parameterizing cortical and subcortical surfaces with respect to simpler algebraic surfaces such as a unit sphere is needed to establish a standard coordinate system. However, polynomial regression type of local parameterization is ill-suited for this purpose. For the global surface parameterization, we can use *surface flattening* [29, 30], which is nonparametric in nature. The surface flattening parametrizes a surface by either solving a partial differential equation or optimizing its variational form.

Deformable surface algorithms naturally provide one-to-one maps from cortical surfaces to a sphere since the algorithm initially starts with a spherical mesh and deforms it to match the tissue boundaries [16]. The deformable surface algorithms usually start with the second level of triangular subdivision of an icosahedron as the initial surface. After several iterations of deformation and triangular subdivision, the resulting cortical surface contains very dense triangle elements. There are many surface flattening techniques such as conformal mapping [30, 31, 32] quasi-isometric mapping [33], area preserving mapping [34], and the Laplace equation method [2].

For many surface flattening methods to work, the starting binary object has to be close to star-shape or convex. For shapes with a more complex structure, the methods may create numerical singularities in mapping to the sphere. Surface flattening can destroy the inherent geometrical structure of the cortical surface due to the metric distortion. Any structural or functional analysis associated with the cortex can be performed without surface flattening if the intrinsic geometric method is used [2].

9.2.3 Spherical Harmonic Representation

The spherical harmonic (SPHARM) representation [2] is a widely used subcortical surface parameterization technique [35, 2, 36, 31, 37, 38]. SPHARM represents the coordinates of mesh vertices as a linear combination of spherical harmonics. SPHARM has been mainly used as a data reduction technique for compressing global shape features into a small number of coefficients. The main global geometric features are encoded in low degree coefficients

---

[2]The SPHARM package is available at
www.stat.wisc.edu/∼mchung/softwares/weighted-SPHARM/weighted-SPHARM.html

while the noise are in high degree spherical harmonics [31]. The method has been used to model various subcortical structures such as ventricles [36], hippocampi [38] and cortical surfaces [35]. The spherical harmonics have a global support. So the resulting spherical harmonic coefficients contain the global shape features and it is not possible to directly obtain local shape information from the coefficients only. However, it is still possible to obtain local shape information by evaluating the representation at each fixed vertex, which gives the smoothed version of the coordinates of surfaces. In this fashion, SPHARM can be viewed as mesh smoothing [35, 39]. In this section, we present a brief introduction of SPHARM within a Hilbert space framework.

Suppose there is a bijective mapping between the cortical surface $\mathcal{M}$ and a unit sphere S^2 obtained through a deformable surface algorithm [2]. Consider the parameterization of S^2 by

$$\mathbf{X}(\theta, \varphi) = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta),$$

with $(\theta, \varphi) \in [0, \pi) \otimes [0, 2\pi)$. The polar angle θ is the angle from the north pole and the azimuthal angle φ is the angle along the horizontal cross-section. Using the bijective mapping, we can parameterize functional data f with respect to the spherical coordinates

$$f(\theta, \varphi) = g(\theta, \varphi) + \epsilon(\theta, \varphi), \tag{9.2}$$

where g is a unknown smooth coordinate function and ϵ is a zero mean random field, possibly Gaussian. The error function ϵ accounts for possible mapping errors. The unknown signal g is then estimated in the finite subspace of $\mathcal{L}^2(S^2)$, the space of square integrable functions in S^2, spanned by spherical harmonics in the least squares fashion [35].

Previous imaging and shape modeling literature have used the complex-valued spherical harmonics [40, 36, 31, 38]. In practice, however, it is sufficient to use only real-valued spherical harmonics [41, 42], which is more convenient in setting up a real-valued stochastic model (9.2). The relationship between the real- and complex-valued spherical harmonics is given in [43, 42]. The complex-valued spherical harmonics can be transformed into real-valued spherical harmonics using an unitary transform.

The spherical harmonic Y_{lm} of degree l and order m is defined as

$$Y_{lm} = \begin{cases} c_{lm} P_l^{|m|}(\cos \theta) \sin(|m|\varphi), & -l \leq m \leq -1, \\ \frac{c_{lm}}{\sqrt{2}} P_l^{|m|}(\cos \theta), & m = 0, \\ c_{lm} P_l^{|m|}(\cos \theta) \cos(|m|\varphi), & 1 \leq m \leq l, \end{cases}$$

where $c_{lm} = \sqrt{\frac{2l+1}{2\pi} \frac{(l-|m|)!}{(l+|m|)!}}$ and P_l^m is the *associated Legendre polynomial* of order m [41, 44], which is given by

$$P_l^m(x) = \frac{(1-x^2)^{m/2}}{2^l l!} \frac{d^{l+m}}{dx^{l+m}} (x^2 - 1)^l, x \in [-1, 1].$$

The first few terms of the spherical harmonics are

$$Y_{00} = \frac{1}{\sqrt{4\pi}}, Y_{1,-1} = \sqrt{\frac{3}{4\pi}} \sin \theta \sin \varphi,$$

$$Y_{1,0} = \sqrt{\frac{3}{4\pi}} \cos \theta, Y_{1,1} = \sqrt{\frac{3}{4\pi}} \sin \theta \cos \varphi.$$

The spherical harmonics are orthonormal with respect to the inner product

$$\langle f_1, f_2 \rangle = \int_{S^2} f_1(\Omega) f_2(\Omega) \, d\mu(\Omega),$$

where $\Omega = (\theta, \varphi)$ and the Lebesgue measure $d\mu(\Omega) = \sin\theta d\theta d\varphi$. The norm is then defined as

$$||f_1|| = \langle f_1, f_1 \rangle^{1/2}. \tag{9.3}$$

The unknown mean function g is estimated by minimizing the integral of the squared residual in $\mathcal{H}_k$, the space spanned by up to k-th degree spherical harmonics:

$$\widehat{g}(\Omega) = \arg\min_{h \in \mathcal{H}_k} \int_{S^2} \left| f(\Omega) - h(\Omega) \right|^2 d\mu(\Omega). \tag{9.4}$$

It can be shown that the minimization is given by

$$\widehat{g}(\Omega) = \sum_{l=0}^{k} \sum_{m=-l}^{l} \langle f, Y_{lm} \rangle Y_{lm}(\Omega), \tag{9.5}$$

the Fourier series expansion. The expansion (9.5) has been referred to as the *spherical harmonic representation* [35, 36, 31, 38, 45]. This technique has been used in representing various brain subcortical structures such as hippocampi [38] and ventricles [36] as well as the whole brain cortical surfaces [35, 31]. By taking each component of Cartesian coordinates of mesh vertices as the functional signal f, surface meshes can be parameterized as a function of θ and φ.

The spherical harmonic coefficients can be estimated in least squares fashion. However, for an extremely large number of vertices and expansions, the least squares method may be difficult to directly invert large matrices. Instead, the *iterative residual fitting (IRF)* algorithm [35] can be used to iteratively estimate the coefficients by partitioning the larger problem into smaller subproblems. The IRF algorithm is similar to the matching pursuit method [46]. The IRF algorithm was developed to avoid the computational burden of inverting a large linear problem while the matching pursuit method was originally developed to compactly decompose a time-frequency signal into a linear combination of a pre-selected pool of basis functions. Although increasing the degree of the representation increases the goodness-of-fit, it also increases the number of estimated coefficients quadratically. So it is necessary to stop the iteration at the specific degree k, where the goodness-of-fit and the number of coefficients balance out. This idea was used in determining the optimal degree of SPHARM [35].

The limitation of SPHARM is that it produces the Gibbs phenomenon, i.e., ringing artifacts, for discontinuous and rapidly changing continuous measurements [35, 47]. The Gibbs phenomenon can be effectively removed by weighting the spherical harmonic coefficients exponentially smaller, which makes the representation smooth out rapidly changing signals. The weighted version of SPHARM is related to isotropic diffusion smoothing [29, 48, 2, 18] as well as the diffusion wavelet transform [35, 39, 49].

9.3 Surface Registration

To construct a test statistic locally at each vertex across different surfaces, one must register the surfaces to a common template surface. Nonlinear cortical surface registration is often performed by minimizing objective functions that measure the global fit of two surfaces while maximizing the smoothness of the deformation in such a way that the cortical gyral or sulcal patterns are matched smoothly [18, 50, 19]. There is also a much simpler way of aligning surfaces using SPHARM representation [35]. Before any sort of nonlinear registration is performed, an affine registration is performed to align and orient the global brain shapes.

9.3.1 Affine Registration

Anatomical objects extracted from 3D medical images are aligned using affine transformations to remove the global size differences. Affine registration requires identifying corresponding landmarks either manually or automatically. The affine transform T of point $p = (p_1, \cdots, p_d)' \in \mathbb{R}^d$ to $q = (q_1, \cdots, q_d)'$ is given by

$$q = Rp + c,$$

where the matrix R corresponds to rotation, scaling and shear while c corresponds to translation. Although the affine transform is not linear, it can be made into a linear form by augmenting the transform. The affine transform can be rewritten as

$$\begin{pmatrix} q \\ 1 \end{pmatrix} = \begin{pmatrix} R & c \\ 0 \cdots 0 & 1 \end{pmatrix} \begin{pmatrix} p \\ 1 \end{pmatrix}. \tag{9.6}$$

Let

$$A = \begin{pmatrix} R & c \\ 0 \cdots 0 & 1 \end{pmatrix}.$$

Trivially, A is linear a linear operator. The matrix A is the most often used form for recording the affine registration.

Let $\mathbf{p}_i$ be the i-th landmark and its corresponding affine transformed points $\mathbf{q}_i$. Then we rewrite (9.6) as

$$\underbrace{\begin{pmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{pmatrix}}_{Q} = \begin{pmatrix} R & c \end{pmatrix} \underbrace{\begin{pmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \\ 1 & \cdots & 1 \end{pmatrix}}_{P}.$$

Then the least squares estimation is given as

$$\begin{pmatrix} \widehat{R} & \widehat{c} \end{pmatrix} = QP'(PP')^{-1}.$$

Then the points $\mathbf{p}_i$ are mapped to $\widehat{R}\mathbf{p}_i + \widehat{c}$, which may not coincide with $\mathbf{q}_i$ in general. In practice, landmarks are automatically identified from T1-weighted MRI.

9.3.2 SPHARM Correspondence

Using SPHARM, it is possible to approximately register surfaces with different mesh topology without any optimization. The crude alignment can be done by coinciding the first order ellipsoid meridian and equator in the SPHARM representation [36, 51]. However, this can be improved. Consider SPHARM representation of surface $H = (h_1, h_2, h_3)$ with spherical angles Ω given by

$$h_i(\Omega) = \sum_{l=0}^{k} \sum_{m=-l}^{l} h_{lm}^i Y_{lm}(\Omega),$$

where (v_1, v_2, v_3) are the coordinates of mesh vertices and SPHARM coefficient $h_{lm}^i = \langle v_i, Y_{lm} \rangle$. Consider another SPHARM representation $J = (j_1, j_2, j_3)$ obtained from mesh coordinate w_i:

$$j_i(\Omega) = \sum_{l=0}^{k} \sum_{m=-l}^{l} w_{lm}^i Y_{lm}(\Omega),$$

where $w_{lm}^i = \langle w_i, Y_{lm} \rangle$. Suppose the surface h_i is deformed to $h_i + d_i$ by the amount of displacement d_i. We wish to find d_i that minimizes the discrepancy between $h_i + d_i$ and
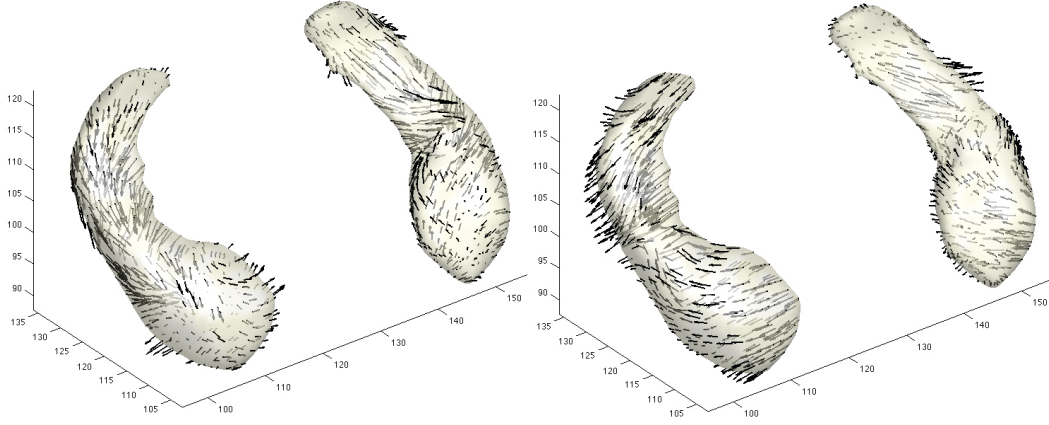
FIGURE 9.2
The displacement vector fields of registering from the hippocampus template to two individual surfaces. The displacement vector field is obtained from the diffeomorphic image registration [1]. The variability in the displacement vector field can be analyzed using a multivariate general linear model [2].

j_i in the subspace $\mathcal{H}_k$ spanned using up to the k-th degree spherical harmonics. It can be shown that

$$\arg \min_{d_i \in \mathcal{H}_k} \left\| \widehat{h}_i + d_i - \widehat{j}_i \right\| = \sum_{l=0}^{k} \sum_{m=-l}^{l} (w_{lm}^i - v_{lm}^i) Y_{lm}(\Omega). \tag{9.7}$$

This implies that the optimal displacement of matching two surfaces is obtained by simply taking the difference between two SPHARM and matching coefficients of the same degree and order. Then a specific point $\widehat{h}_i(\Omega)$ in one surface corresponds to $\widehat{j}_i(\Omega)$ in the other surface. We refer to this point-to-point surface matching as the *SPHARM correspondence* [35]. Unlike other surface registration methods [18, 50, 19], it is not necessary to consider an additional cost function that guarantees the smoothness of the displacement field since the displacement field $d = (d_1, d_2, d_3)$ is already a linear combination of smooth basis functions.

9.3.3 Diffeomorphic Registration

Diffeomorphic image registration is a recently popular technique for registering volume and surface data [1, 52, 53, 54, 55]. From the affine transformed individual surfaces $\mathcal{M}_j$, an additional nonlinear surface registration to the template using the large deformation diffeomorphic metric mapping (LDDMM) framework can be performed [52, 53, 54, 55]. In LDDMM, given the template surface $\mathcal{M}$, the diffeomorphic transformations, which are one-to-one, smooth forward, and inverse transformation, are constructed as follows. We estimate the diffeomorphism d between them as a smooth streamline given by the Lagrangian evolution:

$$\frac{\partial d}{\partial t}(x, t) = v \circ d(x, t)$$

with $d(x, 0) = x$, $t \in [0, 1]$ for time-dependent velocity field v. Note the surfaces $\mathcal{M}_j$ and $\mathcal{M}$ are the start and end points of the diffeomorphism, i.e. $\mathcal{M}_j \circ d(\cdot, 0) = \mathcal{M}_j$ and $\mathcal{M}_j \circ d(\cdot, 1) = \mathcal{M}$. By solving the evolution equation numerically, we obtain the diffeomorphism. By averaging the inverse deformation fields from the template $\mathcal{M}$ to individual subjects,

we may obtain yet another final more refined template. The vector fields v are constrained to be sufficiently smooth to generate diffeomorphic transformations over finite time [56]. Figure 9.2 shows the resulting displacement vector fields of warping from the template to two hippocampal surfaces. In the deformation-based morphometry, variability in the displacement is used to characterize surface growth and differences [2].

9.4   Cortical Surface Features

The human cerebral cortex has the topology of a 2D highly convoluted grey matter shell with an average thickness of 3 mm [2, 16]. The outer boundary of the shell is called the *outer cortical surface* while the inner boundary is called the *inner cortical surface*. Various cortical surface features such as curvatures, local surface area and cortical thickness have been used in quantifying anatomical variations. Among them, cortical thickness has been more often analyzed than other features.

9.4.1   Cortical Thickness

Once we extract both the outer and inner cortical surface meshes, *cortical thickness* can be computed at each mesh vertex. The cortical thickness is defined as the distance between the corresponding vertices between the inner and outer surfaces [16]. There are many different computational approaches in measuring cortical thickness. In one approach, the vertices on the inner triangular mesh are deformed to fit the outer surface by minimizing a cost function that involves bending, stretch and other topological constraints [2]. There is also an alternate method for automatically measuring cortical thickness based on the Laplace equation [57].

The average cortical thickness for each individual is about 3 mm [58]. Cortical thickness varies from 1 to 4 mm depending on the location of the cortex. In normal brain development, it is highly likely that the change of cortical thickness may not be uniform across the cortex. Since different clinical populations are expected to show different patterns of cortical thickness variations, cortical thickness has also been used as a quantitative index for characterizing a clinical population [18]. Cortical thickness varies locally by region and is likely to be influenced by aging, development and disease [59]. By analyzing how cortical thickness varies between clinical and non-clinical populations, we can locate regions on the brain related to a specific pathology.

9.4.2   Surface Area and Curvatures

As in the case of local volume changes in the deformation-based morphometry, the rate of cortical surface area expansion or reduction may not be uniform across the cortical surface [2]. Suppose that cortical surface $\mathcal{M}$ is parameterized by the parameters $\mathbf{u} = (u^1, u^2)$ such that any point $\mathbf{x} \in \mathcal{M}$ can be written as $\mathbf{x} = \mathbf{X}(\mathbf{u})$. Let $\mathbf{X}_i = \partial \mathbf{X}/\partial u^i$ be the partial derivative vectors. The *Riemannian metric tensor* g_{ij} is then given by the inner product between two vectors $\mathbf{X}_i$ and $\mathbf{X}_j$, i.e.

$$g_{ij}(t) = \langle \mathbf{X}_i, \mathbf{X}_j \rangle.$$

The tensor g_{ij} measures the amount of the deviation of the cortical surface from a flat Euclidean plane and can be used to measure lengths, angles and areas on the cortical surface.

Let $g = (g_{ij})$ be a 2×2 matrix of metric tensors. The total surface area of the cortex $\mathcal{M}$ is then given by

$$\int_D \sqrt{\det g} \, d\mathbf{u},$$

where $D = X^{-1}(\mathcal{M})$ is the parameter space [27]. The term $\sqrt{\det g}$ is called the *surface area element* and it measures the transformed area of the unit square in the parameter space D via transformation $X : D \to \mathcal{M}$. The surface area element can be considered as the generalization of the Jacobian determinant, which is used in measuring local volume changes in the tensor-based morphometry [2].

Instead of using the metric tensor formulation, it is possible to quantify local surface area change in terms of the areas of the corresponding triangles in surface meshes. However, this formulation assigns the computed surface area to each face instead of each vertex. This causes a problem in both surface-based smoothing and statistical analysis, where values are required to be given on vertices. Interpolating scalar values on vertices from face values can be done by the weighted average of face values. It is not hard to develop surface-based smoothing and statistical analysis on face values, as a form of dual formulation, but the cortical thickness and the curvature metric are defined on vertices so we end up with two separate approaches: one for metrics defined on vertices and the other for metrics defined on faces. Therefore, the metric tensor approach provides a better unified quantitative framework for the subsequent statistical analysis.

The *principal curvatures* characterize the shape and location of the sulci and gyri, which are the valleys and crests of the cortical surfaces [60, 23, 28]. By measuring the curvature changes, rapidly folding and cortical regions can be localized. The principal curvatures κ_1 and κ_2 can be represented as functions of β_i in quadratic surface (9.1) [25, 27].

9.4.3   Gray Matter Volume

Local volume can be computed using the determinant of the Jacobian of deformation and used in detecting the regions of brain tissue growth and loss in brain development [2]. Compared to the local surface area change, the local volume change measurement is more sensitive to small deformation of the brain. If a unit cube increases its sides by one, the surface area will increase by $2^2 - 1 = 3$ while the volume will increase by $2^3 - 1 = 7$. Therefore, the statistical analysis based on the local volume change is at least twice more sensitive compared to that of the local surface area change. So the local volume change should be able to pick out gray matter tissue growth patterns even when the local surface area change may not.

The gray matter can be considered as a thin shell bounded by two surfaces with varying cortical thickness. In most deformable surface algorithms like FreeSurfer, each triangle on the outer surface has a corresponding triangle on the inner surface. Let $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ be the three vertices of a triangle on the outer surface and $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ be the corresponding three vertices on the inner surface such that $\mathbf{p}_i$ is linked to $\mathbf{q}_i$. Then the volume of the triangular prism is given by the sum of the determinants

$$D(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{q}_1) + D(\mathbf{p}_2, \mathbf{p}_3, \mathbf{q}_1, \mathbf{q}_2) + D(\mathbf{p}_3, \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$$

where

$$D(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}) = |\det(\mathbf{a} - \mathbf{d}, \mathbf{b} - \mathbf{d}, \mathbf{c} - \mathbf{d})|/6$$

is the volume of a tetrahedron whose vertices are $\{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\}$. Afterwards, the total gray matter volume can be estimated by summing the volumes of all triangular prisms [2].

9.5   Surface Data Smoothing

Cortical surface mesh extraction and cortical thickness computation are expected to introduce noise [18, 61, 16]. To counteract this, surface-based data smoothing is necessary [29, 62, 48, 2]. For 3D whole brain volume data, Gaussian kernel smoothing is desirable in many statistical analyses [63, 64]. Gaussian kernel smoothing weights neighboring observations according to their 3D Euclidean distance. Specifically, Gaussian kernel smoothing of functional data or image $f(\mathbf{x}), \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ with *full width at half maximum* $(FWHM) = 4(\ln 2)^{1/2}\sqrt{t}$ is defined as the convolution of the Gaussian kernel with f:

$$F(\mathbf{x}, t) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-(x-y)^2/4t} f(y) dy. \tag{9.8}$$

However, due to the convoluted nature of the cortex, whose geometry is non-Euclidean, we cannot directly use the formulation (9.8) on the cortical surface. For data that lie on a 2D surface, smoothing must be weighted according to the geodesic distance along the surface, which is not straightforward [29, 2, 65].

9.5.1   Diffusion Smoothing

By formulating Gaussian kernel smoothing as a solution of a diffusion equation on a Riemannian manifold, the Gaussian kernel smoothing approach can be generalized to an arbitrary curved surface. This generalization is called *diffusion smoothing* and was first introduced in the analysis of fMRI data on the cortical surface [29] and cortical thickness [65] in 2001.

It can be shown that Gaussian kernel smoothing (9.8) is the integral solution of the n-dimensional diffusion equation

$$\frac{\partial F}{\partial t} = \Delta F \tag{9.9}$$

with the initial condition $F(\mathbf{x}, 0) = f(\mathbf{x})$, where

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$$

is the Laplacian in n-dimensional Euclidean space. Hence the Gaussian kernel smoothing is equivalent to the diffusion of the initial data $f(\mathbf{x})$ after time t.

Diffusion equations have been widely used in image processing as a form of noise reduction starting with Perona and Malik in 1990 [66]. Numerous diffusion techniques have been developed for surface data smoothing [29, 2, 48, 62, 18, 67]. When applying 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* [68, 27]. The approach taken in [29] is based on a local flattening of the cortical surface and estimating the planar Laplacian, which may not be as accurate as the cotan estimation based on the finite element method (FEM) given in [65]. Further, the FEM approach completely avoids the use of surface flattening and parameterization; thus, it is more robust.

For given Riemannian metric tensor g_{ij}, the Laplace-Beltrami operator Δ is defined as

$$\Delta F = \sum_{i,j} \frac{1}{|g|^{1/2}} \frac{\partial}{\partial u^i} \left(|g|^{1/2} g^{ij} \frac{\partial F}{\partial u^j} \right), \tag{9.10}$$
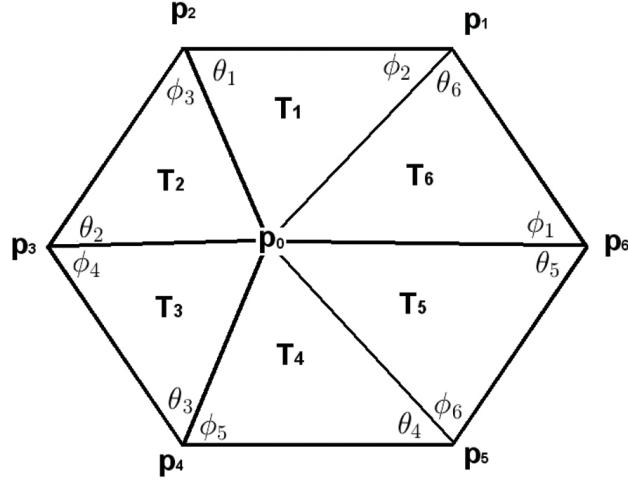
FIGURE 9.3
A typical triangulation in the neighborhood of $\mathbf{p} = \mathbf{p}_0$ in a surface mesh.

where $(g^{ij}) = g^{-1}$ [68]. Note that when g becomes an identity matrix, the Laplace-Beltrami operator reduces to the standard 2D Laplacian:

$$\Delta F = \frac{\partial^2 F}{\partial (u^1)^2} + \frac{\partial^2 F}{\partial (u^2)^2}.$$

Using FEM on the triangular cortical mesh, it is possible to estimate the Laplace-Beltrami operator as the linear weights of neighboring vertices using the cotan formulation, which is first given in [65].

Let $\mathbf{p}_1, \cdots, \mathbf{p}_m$ be m neighboring vertices around the central vertex $\mathbf{p} = \mathbf{p}_0$ (Figure 9.3). Then the estimated Laplace-Beltrami operator is given by

$$\widehat{\Delta F}(\mathbf{p}) = \sum_{i=1}^{m} w_i \big(F(\mathbf{p}_i) - F(\mathbf{p}) \big) \tag{9.11}$$

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 the i-th triangle (Figure 9.3).

FEM estimation (9.11) is an improved formulation from the previous attempt in diffusion smoothing [29], where the Laplacian is simply estimated as the planar Laplacian after locally flattening the triangular mesh consisting of nodes $\mathbf{p}_0, \cdots, \mathbf{p}_m$ onto a flat plane. Afterwards, the finite difference (FD) scheme can be used to iteratively solve the diffusion equation at each vertex $\mathbf{p}$:

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

with the initial condition $F(\mathbf{p}, 0) = f(\mathbf{p})$ and fixed $\delta t = t_{n+1} - t_n$. After N iterations, the FD gives the diffusion of the initial data f after time $N\delta t$. If the diffusion were applied to Euclidean space, it would be approximately equivalent to Gaussian kernel smoothing with

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

For large meshes, computing the linear weights for the Laplace-Beltrami operator takes a fair amount of time, but once the weights are computed, it is applied repeatedly throughout the iterations as a matrix multiplication. Unlike Gaussian kernel smoothing, diffusion smoothing is an iterative procedure.

9.5.2   Iterated Kernel Smoothing

Diffusion smoothing use FEM and FD, which are known to suffer numerical instability if sufficiently small step size is not chosen in the forward Euler scheme. To remedy the problem associated with diffusion smoothing, *iterative kernel smoothing*[3] was introduced [18]. The method has been used in smoothing various cortical surface data: cortical curvatures [69, 70], cortical thickness [71, 72], hippocampus [73, 74], magnetoencephalography (MEG) [75] and functional-MRI [76, 77]. This and its variations are probably the most widely used method for smoothing brain surface data at this moment. In iterated kernel smoothing, kernel weights are spatially adapted to follow the shape of the heat kernel in a discrete fashion.

The *n-th iterated kernel smoothing* of signal $f \in L^2(\mathcal{M})$ with kernel K_σ is defined as

$$K_\sigma^{(n)} * f(p) = \underbrace{K_\sigma * \cdots * K_\sigma}_{n \text{ times}} * f(p),$$

where σ is the bandwidth of the kernel. If K_σ is a heat kernel, we have the following iterative relation [18]:

$$K_\sigma * f(p) = K_{\sigma/n}^{(n)} * f(p). \tag{9.12}$$

The relation (9.12) shows that kernel smoothing with large bandwidth σ can be decomposed into n repeated applications of kernel smoothing with smaller bandwidth σ/n. This idea can be used to approximate the heat kernel. When the bandwidth is small, the heat kernel behaves like the Dirac-delta function and, using the parametrix expansion [78, 79], we can approximate it locally using the Gaussian kernel.

Let $\mathbf{p}_1, \cdots, \mathbf{p}_m$ be m neighboring vertices of vertex $\mathbf{p} = \mathbf{p}_0$ in the mesh. The geodesic distance $d(\mathbf{p}, \mathbf{p}')$ between $\mathbf{p}$ and its adjacent vertex $\mathbf{p}_i$ is the length of edge between these two vertices in the mesh. Then the discretized and normalized heat kernel is given by

$$W_\sigma(\mathbf{p}, \mathbf{p}_i) = \frac{\exp\left(-\frac{d(\mathbf{p}, \mathbf{p}_i)^2}{4\sigma}\right)}{\sum_{j=0}^m \exp\left(-\frac{d(\mathbf{p} - \mathbf{p}_j)^2}{4\sigma}\right)}.$$

Note $\sum_{i=0}^m W_\sigma(\mathbf{p}, \mathbf{p}_i) = 1$. For small bandwidth, all the kernel weights are concentrated near the center, so we only need to worry about the first neighbors of a given vertex in a surface mesh. The discrete version of heat kernel smoothing on a triangle mesh is then given by

$$W_\sigma * f(p) = \sum_{i=0}^m W_\sigma(p, p_i) f(p_i).$$

The discrete kernel smoothing should converge to heat kernel smoothing as the mesh resolution increases. This is the form of the Nadaraya-Watson estimator [80] in statistics. Instead of performing a single kernel smoothing with large bandwidth $n\sigma$, we perform n iterated kernel smoothing with small bandwidth σ as follows $W_\sigma^{(n)} * f(p)$.

---

[3]MATLAB package: `www.stat.wisc.edu/~mchung/softwares/hk/hk.html`.

9.5.3 Heat Kernel Smoothing

The recently proposed *heat kernel smoothing*[4] framework constructs the heat kernel analytically using the eigenfunctions of the Laplace-Beltrami operator [81]. This method avoids the need for the linear approximation used in iterative kernel smoothing that compounds the approximation error at each iteration. The method represents isotropic heat diffusion analytically as a series expansion so it avoids the numerical convergence issues associated with solving the diffusion equations numerically [29, 2, 67]. This technique is different from other diffusion-based smoothing methods in that it bypasses the various numerical problems such as numerical instability, slow convergence, and accumulated linearization error.

Although recently there have been a few studies that introduce heat kernel in computer vision and machine learning [82], they mainly use heat kernel to compute shape descriptors [83, 84] or to define a multi-scale metric [85]. These studies did not use heat kernel in smoothing functional data on manifolds. Further, most kernel methods in machine learning deal with the linear combination of kernels as a solution to penalized regressions, which significantly differs from the heat kernel smoothing framework, which does not have a penalized cost function. There are log-Euclidean and exponential map frameworks on manifolds, where the main interest is in computing the Fréchet mean along the tangent space [86, 87]. Such approaches or related methods in [88], the Nadaya-Watson type of kernel regression is reformulated to learn the shape or image means in a population. In the heat kernel smoothing framework, we are not dealing with manifold data but scalar data defined on a manifold, so there is no need for exploiting the manifold structure of the data itself.

Let Δ be the Laplace-Beltrami operator on $\mathcal{M}$. Solving the eigenvalue equation

$$\Delta\psi_j = -\lambda\psi_j, \tag{9.13}$$

we order eigenvalues

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots,$$

and corresponding eigenfunctions $\psi_0, \psi_1, \psi_2, \cdots$ [78, 18, 89, 90]. The eigenfunctions ψ_j form an orthonormal basis in $L^2(\mathcal{M})$, the space of square integrable functions in $\mathcal{M}$. Figure 9.4 shows the first four LB-eigenfunctions on a hippocampal surface.

There is extensive literature on the use of eigenvalues and eigenfunctions of the Laplace-Beltrami operator in medical imaging and computer vision [89, 91, 92, 93, 94, 95]. The eigenvalues have been used in caudate shape discriminators [96]. Qiu et al. used eigenfunctions in constructing splines on cortical surfaces [91]. Reuter used the topological features of eigenfunctions [93]. Shi et al. used the Reeb graph of the second eigenfunction in shape characterization and landmark detection in cortical and subcortical structures [97, 90]. Lai et al. used the critical points of the second eigenfunction as anatomical landmarks for colon surfaces [98].
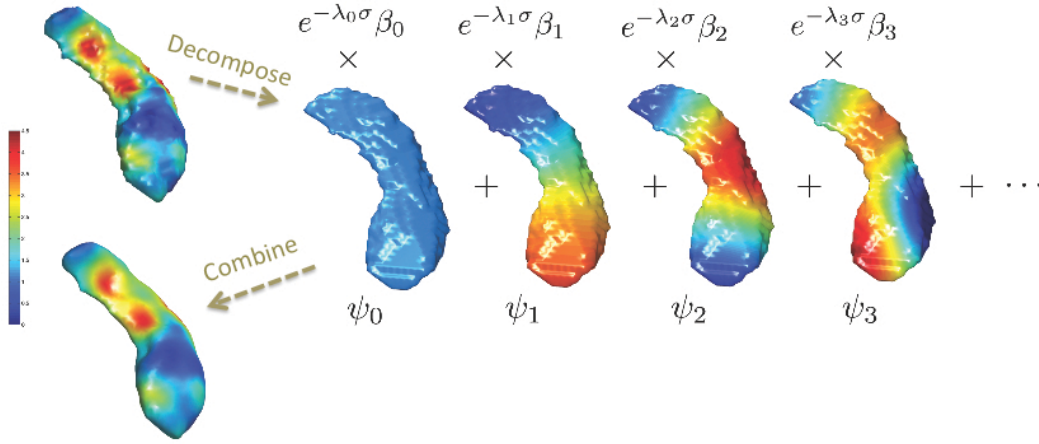
Using the eigenfunctions, *heat kernel* $K_\sigma(p, q)$ is defined as

$$K_\sigma(p, q) = \sum_{j=0}^{\infty} e^{-\lambda_j\sigma}\psi_j(p)\psi_j(q), \tag{9.14}$$

where σ is the bandwidth of the kernel. The heat kernel is the generalization of a Gaussian kernel. Then *heat kernel smoothing* of functional measurement Y is defined as

$$K_\sigma * Y(p) = \sum_{j=0}^{\infty} e^{-\lambda_j\sigma}\beta_j\psi_j(p), \tag{9.15}$$

---

[4]MATLAB package: `brainimaging.waisman.wisc.edu/~chung/lb`.

FIGURE 9.4
Schematic of hat kernel smoothing on a hippocampal surface. Given a noisy functional data on the surface, the Laplace-Beltrami eigenfunctions ψ_j are computed and their exponentially weighted Fourier coefficients $\exp -\lambda_j \sigma$ are multiplied as a form a regression. This process smoothes out the noisy functional signal with bandwidth σ.

where $\beta_j = \langle Y, \psi_j \rangle$ are Fourier coefficients [18]. Kernel smoothing $K_\sigma * Y$ is taken as the estimate for the unknown mean signal θ. The degree for truncating the series expansion can be automatically determined using the forward model selection procedure [35]. Figure 9.4 shows the heat kernel smoothing results with the bandwidth $\sigma = 0.5$ and $k = 500$ number of LB-eigenfunctions.

Unlike previous approaches to heat diffusion [29, 2, 67, 99], heat kernel smoothing avoids the direct numerical discretization of the diffusion equation. Instead, we discretize the basis functions of given manifold $\mathcal{M}$ by solving for the eigensystem (9.13) and obtain λ_j and ψ_j. This provides more robust stable smoothing results compared to diffusion smoothing or iterated kernel smoothing approaches.

9.6 Statistical Inference on Surfaces

Surface measurements such as cortical thickness, curvatures, or fMRI responses can be modeled as random fields on the cortical surface:

$$Y(x) = \mu(x) + \epsilon(x), x \in \mathcal{M}, \tag{9.16}$$

where the deterministic part μ is the unknown mean of the observed functional measurement Y and ϵ is a mean zero random field. The functional measurements on the brain surface is often modeled using the general linear models (GLMs) or its multivariate version. Various statistical models are proposed for estimating and modeling the signal component $\mu(x)$ [23, 2, 35] but the majority of the methods are all based on GLM. GLMs have been implemented in the brain image analysis packages such as SPM and fMRISTAT[5].

---

[5]The fMRISTAT package is available at `www.math.mcgill.ca/keith/fmristat`.

9.6.1 General Linear Models

We set up a GLM at each mesh vertex. Let y_i be the response variable, which is mainly coming from images and $\mathbf{x}_i = (x_{i1}, \cdots, x_{ip})$ to be the variables of interest and $\mathbf{z}_i = (z_{i1}, \cdots, z_{ik})$ to be nuisance variables corresponding to the i-th subject. Assume there are n subjects, i.e., $i = 1, \cdots, n$. We are interested in testing the significance of variables $\mathbf{x}_i$ while accounting for nuisance covariates $\mathbf{z}_i$. Then we set up GLM

$$y_i = \mathbf{z}_i \boldsymbol{\lambda} + \mathbf{x}_i \boldsymbol{\beta} + \epsilon_i$$

where $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_k)'$ and $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_p)'$ are unknown parameter vectors to be estimated. We assume ϵ to be the usual zero mean Gaussian noise, although the distributional assumption is not required for the least squares estimation. We test hypotheses

$$H_0 : \boldsymbol{\beta} = 0 \text{ vs. } H_1 : \boldsymbol{\beta} \neq 0.$$

Subsequently the inference is done by constructing the F-statistic with p and $n - p - k$ degrees of freedom. GLMs have been used in quantifying cortical thickness, for instance, in child development [2, 18] and amygdala shape differences in autism [2].

In the hippocampus case study, the first T1-weighted MRI scans are taken at 11.6 ± 3.7 years for $n = 124$ children using a 3T GE SIGNA scanner. Variables `age` and `gender` are available. We also have variable `income`, which is a binary dummy variable indicating whether the subjects are from high- or low-income families. A total of 124 children and adolescents are from high- ($> 75000\$; n = 86$) and low-income ($< 35000\$, n = 38$) parents respectively. In addition to this cross-sectional data, longitudinal data was available for 82 of these subjects ($n = 66, > 75000\$; n = 16, < 35000\$$). The second MRI scan was acquired for these 82 subjects about 2 years later at 14 ± 3.9 years. For now, we will simply ignore the correlation between the scans within a subject and will treat them as independent.

On the template surface, we have the displacement vector fields of mapping from the template to individual subjects (Figure 9.2). We take the length of the surface displacement, denoted as `deformation`, with respect to the template as the response variable. The displacement measures the shape difference with respect to the template. However, since the length measurement is noisy, surface-based smoothing is necessary. We have used heat kernel smoothing to smooth out noise with the bandwidth 1 and 500 LB-eigenfunctions. Then we set up the GLM:
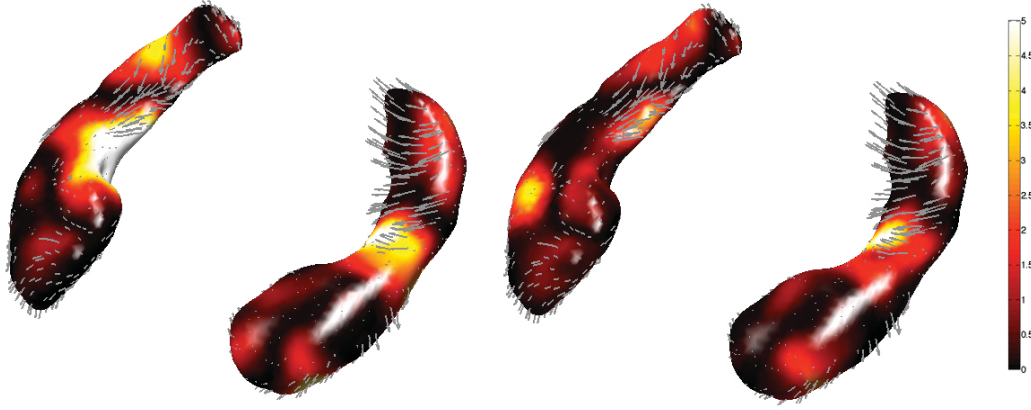
$$\texttt{deformation} = \lambda_1 + \lambda_2 \texttt{age} + \beta_1 \texttt{income} + \epsilon$$

and test for the significance of β_1 at each mesh vertex. Figure 9.5-left shows the F-statistic result on testing β_1.

9.6.2 Multivariate General Linear Models

The multivariate general linear models (MGLMs) have been also used in modeling multivariate imaging features on brain surfaces. These models generalize a widely used univariate GLM by incorporating vector valued responses and explanatory variables [100, 101, 3, 102, 103, 2]. Hotelling's T^2 statistic is a special case of MGLM that has been used primarily for inference on surface shapes, deformations and multi-scale features [104, 2, 105, 106, 107, 108]. An example of this approach is Hotelling's T^2-statistic applied in determining the 3D brain morphology of HIV/AIDS patients [109]. Hotelling's T^2-statistic is also applied to 2D deformation tensor at each mesh vertex on the hippocampal surface as a way to characterize Alzheimers diseases [110].

Suppose there are a total of n subjects and p multivariate features of interest at each

FIGURE 9.5
F-statistics maps on testing the significant hippocampus shape difference on the income level while controlling for age and gender. The arrows show the deformation differences between the groups (high income − low income). The fixed-effects result (left) is obtained by treating the repeat scans as independent. The mixed-effects result (right) is obtained by explicating modeling the covariance structure of the repeat scans with a subject. Both results are not statistically significant under multiple comparisons even at 0.1 level.

voxel. For MGLM to work, n should be significantly larger than p. Let $\mathbf{J}_{n \times p} = (J_{ij})$ be the measurement matrix, where J_{ij} is the measurement for subject i and for the j-th feature. The subscripts denote the dimension of the matrix. All the measurements over subjects for the j-th feature are denoted as $\mathbf{x}_j = (J_{1j}, \cdots, J_{nj})'$. The measurement vector for the i-th subject is denoted as $\mathbf{y}_i = (J_{i1}, \cdots, J_{ip})'$. $\mathbf{y}_i$ is expected to be distributed identically and independently over subjects. Note that

$$\mathbf{J} = (\mathbf{x}_1, \cdots, \mathbf{x}_p) = (\mathbf{y}_1, \cdots, \mathbf{y}_n)'.$$

We may assume the covariance matrix of $\mathbf{y}_i$ to be

$$\mathrm{Cov}(\mathbf{y}_1) = \cdots = \mathrm{Cov}(\mathbf{y}_n) = \mathbf{\Sigma}_{p \times p} = (\sigma_{kl}).$$

With these notations, we now set up the following MGLM at each mesh vertex:

$$\mathbf{J}_{n \times p} = \mathbf{X}_{n \times k}\mathbf{B}_{k \times p} + \mathbf{Z}_{n \times q}\mathbf{G}_{q \times p} + \mathbf{U}_{n \times p}\mathbf{\Sigma}_{p \times p}^{1/2}. \tag{9.17}$$

$\mathbf{X}$ is the matrix of contrasted explanatory variables while $\mathbf{B}$ is the matrix of unknown coefficients. Nuisance covariates are in matrix $\mathbf{Z}$ and the corresponding coefficients are in matrix $\mathbf{G}$. The components of Gaussian random matrix $\mathbf{U}$ are independently distributed with zero mean and unit variance. Symmetric matrix $\mathbf{\Sigma}^{1/2}$ is the square root of the covariance matrix accounting for the spatial dependency across different voxels. In MGLM (9.17), we are usually interested in testing hypotheses

$$H_0 : \mathbf{B} = 0. \ vs. \ H_1 : \mathbf{B} \neq 0.$$

The parameter matrices in the model are estimated via the least squares method. The multivariate test statistics such as Lawley-Hotelling trace or Roy's maximum root are used to test the significance of $\mathbf{B}$. When there is only one voxel, i.e. $p = 1$, these multivariate test statistics collapse to Hotelling's T^2 statistic [102, 2].

9.6.3 Small-n Large-p Problems

GLM are usually fitted in each voxel separately. Instead of fitting GLM at each voxel, one may be tempted to fit the model in the whole brain surface. For FreeSurfer meshes, we need to fit GLM over 300000 vertices, which causes the *small-n large-p problem* [101, 111, 112].

Let $\mathbf{y}_j$ be the measurement vector at the j-th vertex. Assume there are n subjects and total p vertices in the surface. We have the same design matrix $\mathbf{Z}$ for all p vertices. Then we need to estimate the parameter vector $\boldsymbol{\lambda}_j$ in

$$\mathbf{y}_j = \mathbf{Z}\boldsymbol{\lambda}_j \tag{9.18}$$

for each j. Instead of solving (9.18) separately at each vertex, we combine all of them together so that we have matrix equation

$$\underbrace{[\mathbf{y}_1, \cdots, \mathbf{y}_m]}_{\mathbf{Y}} = \mathbf{Z} \underbrace{[\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_m]}_{\boldsymbol{\Lambda}}. \tag{9.19}$$

The least squares estimation of the parameter matrix $\boldsymbol{\Lambda}$ is given by

$$\widehat{\boldsymbol{\Lambda}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}.$$

Note that $\mathbf{Z}$ is of size n by p and $\mathbf{Z}'\mathbf{Z}$ is only invertible when $n \ll p$. The least squares estimation does not provide robust parameter estimates for $n \ll p$, which is the usual case in surface modeling. For small-n large-p problem, GLM need to be regularized using the L1-norm penalty [113, 114, 112, 115, 116, 117].

9.6.4 Longitudinal Models

So far we have only dealt with an imaging data set where the parameters of the model are fixed and do not vary across subjects and scans. Such fixed-effects models are inadequate in modeling the within-subject dependency in longitudinally collected imaging data. However, mixed-effects models can explicitly model such dependency [118, 119, 120, 121]. There are three advantages of the mixed-effects model over the usual fixed-effects model. It explicitly models individual growth patterns and accommodates an unequal number of follow-up image scans per subject and unequal time intervals between scans.

The longitudinal outcome Y_i from the i-th subject is modeled using the mixed-effects model [119] as

$$Y_i = X_i\beta + Z_i\gamma_i + e_i, \tag{9.20}$$

where β is the fixed-effects shared by all subjects. γ_i is the subject-specific random-effects and $e_i \sim N(0, \sigma^2)$ is independent and identically distributed noise. X_i and Z_i are the design matrices corresponding to the fixed and random effects respectively for the i-th subject. We assume $\gamma_i \sim N(0, \Gamma)$ and $\epsilon_i \sim N(0, \Sigma_i)$ with some covariance matrices Γ and Σ_i. Hierarchically we are modeling (9.20) as

$$Y_i | \gamma_i \sim N(X_i\beta + Z_i\gamma_i, \Sigma_i), \ \gamma_i \sim N(0, \Gamma).$$

Γ accounts for covariance among random effect terms. The within-subject variability between the scans is expected to be smaller than between-subject variability and explicitly modeled by Σ_i. The covariance of γ_i and ϵ are expected to have block diagonal structure such that there is no correlation among the scans of different subjects while there is high correlation between the scans of the same subject:

$$\mathbb{V}\left(\begin{array}{c} \gamma_i \\ \epsilon_i \end{array} \right) = \left(\begin{array}{cc} \Gamma & 0 \\ 0 & \Sigma_i \end{array} \right).$$

Subsequently, the overall covariance of Y_i is given by

$$\mathbb{V}Y_i = Z_i \Gamma Z_i' + \Sigma_i.$$

The random-effect contribution is $Z_i \Gamma Z_i'$ while the within-subject contribution is Σ_i.

The parameters and the covariance matrices can be estimated via the restricted maximum likelihood (REML) method [118, 121]. The most widely used tools for fitting the mixed-effects model are the `nlme` library in the `R` statistical package [121]. However, there is no need to use `R` to fit the mixed-effects model. Keith Worsley has implemented the REML procedure in the `SurfStat` package[6] [2, 122].

Here we briefly explain how to set up a longitudinal mixed-effect model in practice. In the usual fixed-effect model, we have a linear model containing the fixed-effect term $\mathtt{age}_i$ for the i-th subject:

$$y_i = \beta_0 + \beta_1 \mathtt{age}_i + \epsilon_i, \tag{9.21}$$

where ϵ_i is assumed to follow independent Gaussian. In (9.21), every subject has identical growth trajectory $\beta_0 + \beta_1 \mathtt{age}$, which is unrealistic. Biologically, each subject is expected to have its own unique growth trajectory. So we assume each subject to have its own intercept $\beta_0 + \gamma_{i0}$ and slope $\beta_1 + \gamma_{i1}$:

$$y_i = \beta_0 + \gamma_{i0} + (\beta_1 + \gamma_{i1})\mathtt{age}_i + \epsilon_i. \tag{9.22}$$

It is reasonable to assume the random vector $\gamma = (\gamma_{i0}, \gamma_{i1})$ to be multivariate normal. The model (9.22) can be decomposed into fixed- and random-effect terms:

$$y_i = (\beta_0 + \beta_1 \mathtt{age}_i) + (\gamma_{i0} + \gamma_{i1}\mathtt{age}_i) + \epsilon_i. \tag{9.23}$$

The fixed-effect term $\beta_0 + \beta_1 \mathtt{age}_i$ models the linear growth of the population while the random-effect term $\gamma_{i0} + \gamma_{i1}\mathtt{age}_i$ models the subject specific growth variations. Incorporating additional factors and interaction terms are done similarly.

In the hippocampus case study, the first MRI scans are taken at 11.6 ± 3.7 years while the second scans are taken at 14 ± 3.9 years. We are interested in determining the effects of income level on the shape of the hippocampus. In Section 9.6.1, we treated the second scans as if they came from independent subjects and modeled them using the fixed-effects model. Now we explicitly incorporate the dependency of repeated scans of the same subject. It is necessary to explicitly model the within-subject variability that is expected to be smaller than between-subject variability. This can be done by introducing a random-effect term. The resulting F-statistic maps are given in Figure 9.5-right. However, we did not detect any region that is affected by income. Thus, we tested the age and income interaction and found the regions of strong interaction (Figure 9.6).

9.6.5 Random Field Theory

Since we need to set up a GLM on every mesh vertex, it becomes a *multiple comparisons* problem. Correcting for multiple comparisons is crucial in determining overall statistical significance in correlated test statistics over the whole surface. For surface data, various methods are proposed: Bonferroni correction, random field theory [123, 3], false discovery rates (FDR) [124, 125, 126], and permutation tests [127]. Among many techniques, the

---

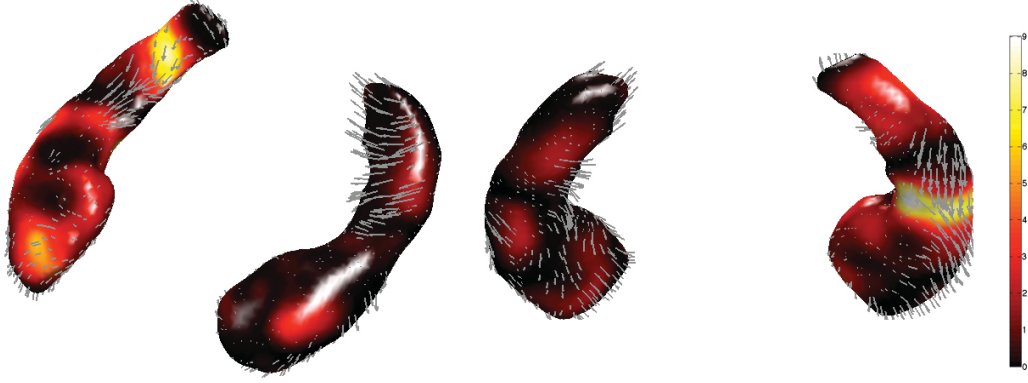[6]The MATLAB package is available at `www.math.mcgill.ca/keith/surfstat`.

FIGURE 9.6
F-statistics maps on testing the interaction between the income level and age while controlling for gender in a linear mixed-effects model. The arrows show the deformation differences between the groups (high income - low income). Significant regions are only found in the tail and midbody regions of the right hippocampus.

random field theory is probably the most natural in relation to surface data smoothing since it is able to explicitly control the amount of smoothing.

The generalization of a continuous stochastic process in $\mathbb{R}^n$ to a higher dimensional abstract space is called a *random field* [128, 2, 63, 106, 129]. In the random field theory [123, 3], measurement Y at position $x \in \mathcal{M}$ is modeled as

$$Y(x) = \mu(x) + \epsilon(x),$$

where μ is the unknown signal to be estimated and ϵ is the measurement error. The measurement error at each fixed x can be modeled as a random variable. Then the collection of random variables $\{\epsilon(x) : x \in \mathcal{M}\}$ is called a *random field* or stochastic process. A measure-theoretic definition is given in [128].

Detecting the regions of statistically significant $\lambda(x)$ for some $x \in \mathcal{M}$ can be done via thresholding the maximum of a random field defined on the cortical surface [3, 4]. For instance, T random field on the surface $\mathcal{M}$ is defined as

$$T(x) = \sqrt{n}\frac{M(x)}{S(x)},$$

where M and S are the sample mean and standard deviation over the n subjects. Under the null hypothesis
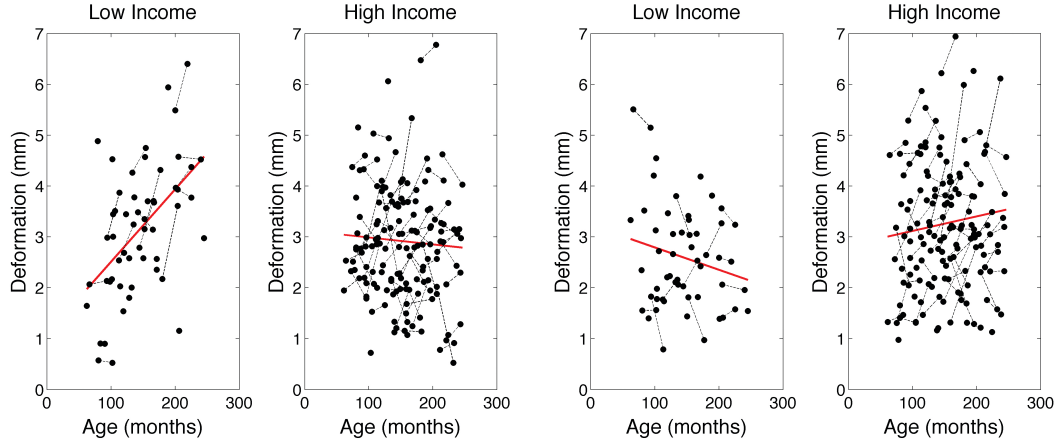
$$H_0 : \mu(x) = 0 \text{ for all } x \in \mathcal{M},$$

$T(x)$ is distributed as a student's T with $n - 1$ degrees of freedom at each voxel x. The pvalue of the local maxima of the T field will give a conservative threshold compared to FDR [3].

For sufficiently high threshold y, we can show that

$$P\left(\max_{x \in \mathcal{M}} T(\mathbf{x}) \geq y\right) \approx \sum_{i=0}^{3} \phi_i(\mathcal{M})\rho_i(y), \tag{9.24}$$

where ρ_i is the i-dimensional EC-density and the Minkowski functional ϕ_i are

$$\phi_0(\mathcal{M}) = 2, \ \phi_1(\mathcal{M}) = 0, \ \phi_2(\mathcal{M} = \|\mathcal{M}\|, \ \phi_3(\mathcal{M}) = 0$$

FIGURE 9.7
The plots showing income level dependent growth differences in the posterior (left) and midbody (right) regions of the right hippocampus. The red lines are the linear regression lines. Scans within a subject are identified by dotted lines.

and $\|\mathcal{M}\|$ is the total surface area of $\mathcal{M}$ (Worsley, 1996a). When diffusion or heat kernel smoothing with given FWHM is applied on surface $\mathcal{M}$, the 0-dimensional and 2-dimensional EC-density becomes

$$\rho_0(y) = \int_y^\infty \frac{\Gamma(\frac{n}{2})}{((n-1)\pi)^{1/2}\Gamma(\frac{n-1}{2})}\Big(1 + \frac{y^2}{n-1}\Big)^{-n/2}\, dy,$$

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

The excursion probability, which is the probability of obtaining false positives for the one-sided alternate hypothesis, is approximated by the following formula:

$$P\Big(\max_{x\in\mathcal{M}} T(x) \geq y\Big) \approx 2\rho_0(y) + \|\mathcal{M}\|\rho_2(y).$$

For smoothing cortical thickness [2, 18], an FWHM of between 20 to 30 mm is recommended. This FWHM reflects the spatial frequency associated with the sulcal pattern. For measurements on smaller subcortical structures such as hippocampus and amygdala, significantly smaller FWHM is recommended. For amygdala and hippocampus, 0.5–1 mm would be sufficient.

In the hippocampus case study, we did not detect any statistically significant group difference at 0.01 level after correcting for multiple comparisons in both the left and right hippocampi. However, we obtained highly focalized regions of group difference in the growth rate, the interaction between income level and age, in the right hippocampus (corrected pvalue = 0.03). The posterior region is enlarging while the midbody and the anterior parts are shrinking in children from low-income families (Figure 9.6 and 9.7). On the other hand, the pattern is opposite for children from high-income families. Note that the right hippocampus is involved in the active maintenance of associations with spatial information [130]. Future studies investigating the relationship between family socioeconomic status and spatial information processing measures are warranted.

Bibliography

[1] B.B. Avants, C.L. Epstein, M. Grossman, and J.C. Gee. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12:26–41, 2008.

[2] M.K. Chung. *Computational Neuroanatomy: The Methods*. World Scientific, 2013.

[3] 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.

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

[5] M.K. Chung, J.L. Hanson, R.J. Davidson, and S.D. Pollak. Effect of family income on hippocampus growth: Longitudinal study. *17th Annual Meeting of the Organization for Human Brain Mapping*, (2697), 2011.

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

[7] M.K. Chung, K.J. Worsley, S. Robbins, T. Paus, J. Taylor, J.N. Giedd, J.L. Rapoport, and A.C. Evans. Deformation-based surface morphometry applied to gray matter deformation. *NeuroImage*, 18:198–213, 2003.

[8] 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.

[9] J.G. Sled, A.P. Zijdenbos, and A.C. Evans. A nonparametric method for automatic correction of intensity nonuniformity in mri data. *IEEE Transactions on Medical Imaging*, 17:87–97, 1988.

[10] K. Kollakian. Performance analysis of automatic techniques for tissue classification in magnetic resonance images of the human brain. Technical Report Master's thesis, Concordia University, Montreal, Quebec, Canada, 1996.

[11] M. Ozkan, B.M. Dawant, and R.J. Maciunas. Neural-network-based segmentation of multi-modal medical images: a comparative and prospective study. *IEEE Transactions on Medical Imaging*, 12:534–544, 1993.

[12] C.D. Good, I.S. Johnsrude, J. Ashburner, R.N.A. Henson, K.J. Friston, and R.S.J. Frackowiak. A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*, 14:21–36, 2001.

[13] W.E. Lorensen and H.E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. In *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, pages 163–169, 1987.

[14] J.A. Sethian. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision and Material Science*. Cambridge University Press, 2002.

[15] C. Davatzikos and R.N. Bryan. Using a deformable surface model to obtain a shape representation of the cortex. *Proceedings of the IEEE International Conference on Computer Vision*, 9:2122–2127, 1995.

[16] 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.

[17] B.D. Rusch, H.C. Abercrombie, T.R. Oakes, S.M. Schaefer, and R.J. Davidson. Hippocampal morphometry in depressed patients and control subjects: Relations to anxiety symptoms. *Biological Psychiatry*, 50:960–964, 2001.

[18] M.K. Chung, S. Robbins, and A.C. Evans. Unified statistical approach to cortical thickness analysis. *Information Processing in Medical Imaging (IPMI), Lecture Notes in Computer Science*, 3565:627–638, 2005.

[19] 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.

[20] C. Davatzikos. Spatial transformation and registration of brain images using elastically deformable models. *Comput. Vis. Image Understanding*, 66:207–222, 1997.

[21] M.I. Miller, A. Banerjee, G.E. Christensen, S.C. Joshi, N. Khaneja, U. Grenander, and L. Matejic. Statistical methods in computational anatomy. *Statistical Methods in Medical Research*, 6:267–299, 1997.

[22] B. Fischl, M.I. Sereno, R. Tootell, and A.M. Dale. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapping*, 8:272–284, 1999.

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

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

[25] W.M. Boothby. *An Introduction to Differential Manifolds and Riemannian Geometry*. Academic Press, London, 2nd edition, 1986.

[26] M.P. do Carmo. *Riemannian Geometry*. Prentice-Hall, Inc., 1992.

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

[28] N. Khaneja, M.I. Miller, and U. Grenander. Dynamic programming generation of curves on brain surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:1260–1265, 1998.

[29] 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.

[30] 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.

[31] X. Gu, Y.L. Wang, T.F. Chan, T.M. Thompson, and S.T. Yau. Genus zero surface conformal mapping and its application to brain surface mapping. *IEEE Transactions on Medical Imaging*, 23:1–10, 2004.

[32] M. K. Hurdal and K. Stephenson. Cortical cartography using the discrete conformal approach of circle packings. *NeuroImage*, 23:S119–S128, 2004.

[33] B. Timsari and R. Leahy. An optimization method for creating semi-isometric flat maps of the cerebral cortex. In *The Proceedings of SPIE, Medical Imaging*, 2000.

[34] C. Brechbuhler, G. Gerig, and O. Kubler. Parametrization of closed surfaces for 3d shape description. *Computer Vision and Image Understanding*, 61:154–170, 1995.

[35] M.K. Chung, R. Hartley, K.M. Dalton, and R.J. Davidson. Encoding cortical surface by spherical harmonics. *Statistica Sinica*, 18:1269–1291, 2008.

[36] G. Gerig, M. Styner, D. Jones, D. Weinberger, and J. Lieberman. Shape analysis of brain ventricles using SPHARM. In *MMBIA*, pages 171–178, 2001.

[37] A. Kelemen, G. Szekely, and G. Gerig. Elastic model-based segmentation of 3-d neuroradiological data sets. *IEEE Transactions on Medical Imaging*, 18:828–839, 1999.

[38] L. Shen, J. Ford, F. Makedon, and A. Saykin. Surface-based approach for classification of 3d neuroanatomical structures. *Intelligent Data Analysis*, 8:519–542, 2004.

[39] M.K. Chung, S.M. Schaefer, C. M. van Reekum, L.P. Schmitz, M. Sutterer, and R.J. Davidson. A unified kernel regression on manifolds detects aging-related changes in the amygdala and hippocampus. *MICCAI, Lecture Notes in Computer Science (LNCS)*, 8674:789–796, 2014.

[40] T. Bulow. Spherical diffusion for 3D surface smoothing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26:1650–1654, 2004.

[41] R. Courant and D. Hilbert. *Methods of Mathematical Physics.* Interscience, New York, English edition, 1953.

[42] H.H.H. Homeier and E.O. Steinborn. Some properties of the coupling coefficients of real spherical harmonics and their relation to gaunt coefficients. *Journal of Molecular Structure: THEOCHEM*, 368:31–37, 1996.

[43] M.A. Blanco, M. Florez, and M. Bermejo. Evaluation of the rotation matrices in the basis of real spherical harmonics. *Journal of Molecular Structure: THEOCHEM*, 419:19–27, 1997.

[44] G. Wahba. *Spline Models for Observational Data.* SIAM, New York, 1990.

[45] L. Shen and M.K. Chung. Large-scale modeling of parametric surfaces using spherical harmonics. In *Third International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)*, 2006.

[46] S. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41:3397–3415, 1993.

[47] A. Gelb. The resolution of the Gibbs phenomenon for spherical harmonics. *Mathematics of Computation*, 66:699–717, 1997.

[48] A. Cachia, J.-F. Mangin, D. Riviére, F. Kherif, N. Boddaert, A. Andrade, D. Papadopoulos-Orfanos, J.-B. Poline, I. Bloch, M. Zilbovicius, P. Sonigo, F. Brunelle, and J. Régis. A primal sketch of the cortex mean curvature: A morphogenesis based approach to study the variability of the folding patterns. *IEEE Transactions on Medical Imaging*, 22:754–765, 2003.

[49] A.P. Hosseinbor, W.H. Kim, N. Adluru, A. Acharya, H.K. Vorperian, and M.K. Chung. The 4d hyperspherical diffusion wavelet: A new method for the detection of localized anatomical variation. *MICCAI, Lecture Notes in Computer Science (LNCS)*, 8675:65–72, 2014.

[50] S.M. Robbins. *Anatomical standardization of the human brain in Euclidean 3-Space and on the cortical 2-manifold.* PhD thesis, School of Computer Science, McGill University, Montreal, Quebec, Canada, 2003.

[51] M. Styner, I. Oguz, S. Xu, C. Brechbuhler, D. Pantazis, J. Levitt, M. Shenton, and G. Gerig. Framework for the statistical shape analysis of brain structures using SPHARM-PDM. In *Insight Journal, Special Edition on the Open Science Workshop at MICCAI*, 2006.

[52] M.I. Miller and A. Qiu. The emerging discipline of computational functional anatomy. *Neuroimage*, 45(1S):S16–S39, 2009.

[53] M. Vaillant, A. Qiu, J. Glaunès, and M.I. Miller. Diffeomorphic metric surface mapping in subregion of the superior temporal gyrus. *NeuroImage*, 34:1149–1159, 2007.

[54] A. Qiu and M.I. Miller. Multi-structure network shape analysis via normal surface momentum maps. *NeuroImage*, 42:1430–1438, 2008.

[55] X. Yang, A. Goh, and A. Qiu. Locally linear diffeomorphic metric embedding (LLDME) for surface-based anatomical shape modeling. *NeuroImage*, 56:149–161, 2011.

[56] P. Dupuis, U. Grenander, and M. I. Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Math.*, 56:587–600, 1998.

[57] 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.

[58] C.C. Henery and T.M. Mayhew. The cerebrum and cerebellum of the fixed human brain: Efficient and unbiased estimates of volumes and cortical surface areas. *Journal of Anatomy*, 167:167–180, 1989.

[59] P. Barta, M.I. Miller, and A. Qiu. A stochastic model for studying the laminar structure of cortex from mri. *IEEE Transactions on Medical Imaging*, 24:728–742, 2005.

[60] A. Bartesaghi and G. Sapiro. A system for the generation of curves on 3d brain images. *Human Brain Mapping*, 14(1):1–15, 2001.

[61] B. Fischl and A.M. Dale. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences (PNAS)*, 97:11050–11055, 2000.

[62] A. Cachia, J.-F. Mangin, D. Riviére, D. Papadopoulos-Orfanos, F. Kherif, I. Bloch, and J. Régis. A generic framework for parcellation of the cortical surface into gyri using geodesic Voronoï diagrams. *Image Analysis*, 7:403–416, 2003.

258

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

[64] A. Rosenfeld and A.C. Kak. *Digital Picture Processing.* Academic Press, New York, 1982.

[65] M.K. Chung, K.J. Worsley, J. Taylor, J. Ramsay, S. Robbins, and A.C. Evans. Diffusion smoothing on the cortical surface. *NeuroImage*, 13:S95, 2001.

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

[67] A.A. Joshi, D. W. Shattuck, P. M. Thompson, and R. M. Leahy. A parameterization-based numerical method for isotropic and anisotropic diffusion smoothing on non-flat surfaces. *IEEE Transactions on Image Processing*, 18:1358–1365, 2009.

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

[69] E. Luders, P.M. Thompson, K.L. Narr, A.W. Toga, L. Jancke, and C. Gaser. A curvature-based approach to estimate local gyrification on the cortical surface. *NeuroImage*, 29:1224–1230, 2006.

[70] C. Gaser, E. Luders, P.M. Thompson, A.D. Lee, R.A. Dutton, J.A. Geaga, K.M. Hayashi, U. Bellugi, A.M. Galaburda, J.R. Korenberg, D.L. Mills, A.W. Toga, and A.L. Reiss. Increased local gyrification mapped in Williams syndrome. *NeuroImage*, 33:46–54, 2006.

[71] E. Luders, K.L. Narr, P.M. Thompson, D.E. Rex, R.P. Woods, H. DeLuca, L. Jancke, and A.W. Toga. Gender effects on cortical thickness and the influence of scaling. *Human Brain Mapping*, 27:314–324, 2006.

[72] J. Bernal-Rusiel, M. Atienza, and J. Cantero. Detection of focal changes in human cortical thickness: Spherical wavelets versus gaussian smoothing. *NeuroImage*, 41:1278–1292, 2008.

[73] L. Shen, A. Saykin, M.K. Chung, H. Huang, J. Ford, F. Makedon, T.L. McHugh, and C.H. Rhodes. Morphometric analysis of genetic variation in hippocampal shape in mild cognitive impairment: Role of an il-6 promoter polymorphism. In *Life Science Society Computational Systems Bioinformatics Conference*, 2006.

[74] H. Zhu, J.G. Ibrahim, N. Tang, D.B. Rowe, X. Hao, R. Bansal, and B.S. Peterson. A statistical analysis of brain morphology using wild bootstrapping. *IEEE Transactions on Medical Imaging*, 26:954–966, 2007.

[75] J. Han, J.S. Kim, C.K. Chung, and K.S. Park. Evaluation of smoothing in an iterative lp-norm minimization algorithm for surface-based source localization of meg. *Physics in Medicine and Biology*, 52:4791–4803, 2007.

[76] D.J. Hagler Jr., A.P. Saygin, and M.I. Sereno. Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. *NeuroImage*, 33:1093–1103, 2006.

[77] H.J. Jo, J.-M. Lee, J.-H. Kim, Y.-W. Shin, I.-Y. Kim, J.S. Kwon, and S.I. Kim. Spatial accuracy of fmri activation influenced by volume- and surface-based spatial smoothing techniques. *NeuroImage*, 34:550–564, 2007.

[78] S. Rosenberg. *The Laplacian on a Riemannian Manifold.* Cambridge University Press, 1997.

[79] F.-Y. Wang. Sharp explicit lower bounds of heat kernels. *Annals of Probability*, 24:1995–2006, 1997.

[80] P. Chaudhuri and J.S. Marron. Scale space view of curve estimation. *The Annals of Statistics*, 28:408–428, 2000.

[81] S. Seo, M.K. Chung, and H.K. Vorperian. Heat kernel smoothing using Laplace-Beltrami eigenfunctions. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2010*, volume 6363 of *Lecture Notes in Computer Science*, pages 505–512, 2010.

[82] Mikhail Belkin, Partha Niyogi, and Vikas Sindhwani. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *The Journal of Machine Learning Research*, 7:2399–2434, 2006.

[83] J. Sun, M. Ovsjanikov, and L. J. Guibas. A concise and provably informative multi-scale signature based on heat diffusion. *Comput. Graph. Forum*, 28:1383–1392, 2009.

[84] M. M. Bronstein and I. Kokkinos. Scale-invariant heat kernel signatures for non-rigid shape recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1704–1711, 2010.

[85] F. de Goes, S. Goldenstein, and L. Velho. A hierarchical segmentation of articulated bodies. *Computer Graphics Forum*, 27:1349–1356, 2008.

[86] B.C. Davis, P.T. Fletcher, E. Bullitt, and S. Joshi. Population shape regression from random design data. *International Journal of Computer Vision*, 90:255–266, 2010.

[87] P.T. Fletcher, C. Lu, S.M. Pizer, and S. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, 23:995–1005, 2004.

[88] Samuel Gerber, Tolga Tasdizen, and Ross Whitaker. Dimensionality reduction and principal surfaces via kernel map manifolds. In *IEEE 12th International Conference on Computer Vision (ICCV)*, pages 529–536, 2009.

[89] B. Lévy. Laplace-Beltrami eigenfunctions towards an algorithm that "understands" geometry. In *IEEE International Conference on Shape Modeling and Applications*, page 13, 2006.

[90] Y. Shi, I. Dinov, and A. W. Toga. Cortical shape analysis in the Laplace-Beltrami feature space. In *12th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2009)*, volume 5762 of *Lecture Notes in Computer Science (LNCS)*, pages 208–215, 2009.

[91] A. Qiu, D. Bitouk, and M.I. Miller. Smooth functional and structural maps on the neocortex via orthonormal bases of the Laplace-Beltrami operator. *IEEE Transactions on Medical Imaging*, 25:1296–1396, 2006.

[92] M. Reuter, F.-E. Wolter, M. Shenton, and M. Niethammer. Laplace-Beltrami eigenvalues and topological features of eigenfunctions for statistical shape analysis. *Computer-Aided Design*, 41:739–755, 2009.

[93] M. Reuter. Hierarchical shape segmentation and registration via topological features of Laplace-Beltrami eigenfunctions. *International Journal of Computer Vision*, 89:287–308, 2010.

[94] H. Zhang, O. van Kaick, and R. Dyer. Spectral methods for mesh processing and analysis. In *EUROGRAPHICS*, pages 1–22, 2007.

[95] H. Zhang, O. van Kaick, and R. Dyer. Spectral mesh processing. *Computer Graphics Forum*, 29:1865–1894, 2010.

[96] M. Niethammer, M. Reuter, F. Wolter, S. Bouix, N. Peinecke, M. Koo, and M.E. Shenton. Global medical shape analysis using the Laplace-Beltrami spectrum. *Lecture Notes in Computer Science*, 4791:850, 2007.

[97] Y. Shi, R. Lai, K. Kern, N. Sicotte, I. Dinov, and A. W. Toga. Harmonic surface mapping with Laplace-Beltrami eigenmaps. In *11th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2008)*, volume 5242 of *Lecture Notes in Computer Science (LNCS)*, pages 147–154, 2008.

[98] Z. Lai, J. Hu, C. Liu, V. Taimouri, D. Pai, J. Zhu, J. Xu, and J. Hua. Intra-patient supine-prone colon registration in CT colonography using shape spectrum. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2010*, volume 6361 of *Lecture Notes in Computer Science*, pages 332–339, 2010.

[99] T. Tasdizen, R. Whitaker, P. Burchard, and S. Osher. Geometric surface smoothing via anisotropic diffusion of normals. In *Geometric Modeling and Processing*, pages 687–693, 2006.

[100] T.W. Anderson. *An Introduction to Multivariate Statistical Analysis*. Wiley, 2nd edition, 1984.

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

[102] K.J. Worsley, J.E. Taylor, F. Tomaiuolo, and J. Lerch. Unified univariate and multivariate random field theory. *NeuroImage*, 23:S189–195, 2004.

[103] J.E. Taylor and K.J. Worsley. Random fields of multivariate test statistics, with applications to shape analysis. *Annals of Statistics*, 36:1–27, 2008.

[104] J. Cao and K.J. Worsley. The detection of local shape changes via the geometry of Hotelling's T2 fields. *Annals of Statistics*, 27:925–942, 1999.

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

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

[107] W.H. Kim, D. Pachauri, C. Hatt, M.K. Chung, S. Johnson, and V. Singh. Wavelet based multi-scale shape features on arbitrary surfaces for cortical thickness discrimination. In *Advances in Neural Information Processing Systems*, pages 1250–1258, 2012.

[108] 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.

[109] N. Lepore, C.A. Brun, M.C. Chiang, Y.Y. Chou, R.A. Dutton, K.M. Hayashi, O.L. Lopez, H.J. Aizenstein, A.W. Toga, J.T. Becker, and P.M. Thompson. Multivariate statistics of the Jacobian matrices in tensor based morphometry and their application to HIV/AIDS. *Lecture Notes in Computer Science*, pages 191–198, 2006.

[110] Y. Wang, Y. Song, P. Rajagopalan, T. An, K. Liu, Y.Y. Chou, B. Gutman, A.W. Toga, P.M. Thompson, and the Alzheimer's Disease Neuroimaging Initiative. Surface-based TBM boosts power to detect disease effects on the brain: An N= 804 ADNI study. *NeuroImage*, 56:1993–2010, 2011.

[111] J. Schäfer and K. Strimmer. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*, 4:32, 2005.

[112] M.K. Chung. *Statistical and Computational Methods in Brain Image Analysis*. CRC Press, 2013.

[113] O. Banerjee, L.E. Ghaoui, A. d'Aspremont, and G. Natsoulis. Convex optimization techniques for fitting sparse Gaussian graphical models. In *Proceedings of the 23rd International Conference on Machine Learning*, page 96, 2006.

[114] O. Banerjee, L. El Ghaoui, and A. d'Aspremont. Model selection through sparse maximum likelihood estimation for multivariate Gaussian or binary data. *The Journal of Machine Learning Research*, 9:485–516, 2008.

[115] J. Friedman, T. Hastie, and R. Tibshirani. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9:432, 2008.

[116] S. Huang, J. Li, L. Sun, J. Ye, A. Fleisher, T. Wu, K. Chen, and E. Reiman. Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. *NeuroImage*, 50:935–949, 2010.

[117] R. Mazumder and T. Hastie. Exact covariance thresholding into connected components for large-scale graphical LASSO. *The Journal of Machine Learning Research*, 13:781–794, 2012.

[118] J. Fox. *An R and S-Plus Companion to Applied Regression*. Sage Publications, Inc., 2002.

[119] J.K. Milliken and S.D. Edland. Mixed effect models of longitudinal Alzheimer's disease data: A cautionary note. *Statist. Med.*, 19:1617–1629, 2000.

[120] G. Molenberghs and G. Verbeke. *Models for Discrete Longitudinal Data*. Springer, 2005.

[121] J.C. Pinehiro and D.M. Bates. *Mixed Effects Models in S and S-Plus*. Springer, 3rd edition, 2002.

[122] K.J. Worsley, J.E. Taylor, F. Carbonell, M.K. Chung, E. Duerden, B. Bernhardt, O. Lyttelton, M. Boucher, and A.C. Evans. SurfStat: A MATLAB toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. *NeuroImage*, 47:S102, 2009.

[123] 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.

262

[124] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of Royal Statistical Society, Series. B*, 57:289–300, 1995.

[125] Y. Benjamini and D. Yekutieli. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, pages 1165–1188, 2001.

[126] C.R. Genovese, N.A. Lazar, and T. Nichols. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15:870–878, 2002.

[127] T.E. Nichols and A.P. Holmes. Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, 15:1–25, 2002.

[128] R.J. Adler and J.E. Taylor. *Random Fields and Geometry*. Springer Verlag, 2007.

[129] A.M. Yaglom. *Correlation Theory of Stationary and Related Random Functions Vol. I: Basic Results*. Springer-Verlag, 1987.

[130] C. Piekema, R.P.C. Kessels, R.B. Mars, K.M. Petersson, and G. Fernández. The right hippocampus participates in short-term memory maintenance of object-location associations. *NeuroImage*, 33:374–382, 2006.