Unified Heat Kernel Regression for Diffusion, Kernel Smoothing and Wavelets on Manifolds and Its Application to Mandible Growth Modeling in CT Images

Moo K. Chung[a], Anqi Qiu[b], Seongho Seo[c], Houri K. Vorperian[a]

[a] *University of Wisconsin, Madison*
[b] *National University of Singapore*
[c] *Seoul National University, Korea*

`mkchung@wisc.edu`

Abstract

We present a novel kernel regression framework for smoothing scalar surface data using the Laplace-Beltrami eigenfunctions. Starting with the heat kernel constructed from the eigenfunctions, we formulate a new bivariate kernel regression framework as a weighted eigenfunction expansion with the heat kernel as the weights. The new kernel regression is mathematically equivalent to isotropic heat diffusion, kernel smoothing and recently popular diffusion wavelets. Unlike many previous partial differential equation based approaches involving diffusion, our approach represents the solution of diffusion analytically reducing numerical inaccuracy and slow convergence. The numerical implementation is validated on a unit sphere using spherical harmonics. As an illustration, we have applied the method in characterizing the localized growth pattern of mandible surfaces obtained in CT images between ages 0 and 20 by regressing the length of displacement vectors with respect to the template surface.

Keywords: Heat kernel smoothing, Laplace-Beltrami eigenfunctions, Mandible growth, Surface-based morphometry, Diffusion wavelet

1. Introduction

In medical imaging, anatomical surfaces extracted from MRI and CT are often represented as triangular meshes. Image segmentation and surface extraction process themselves are likely to introduce noise to the mesh coordinates. It is imperative to reduce the mesh noise while preserving the geometric details of the anatomical structures for various applications.

Diffusion equations have been widely used in image processing as a form of noise reduction starting with Perona and Malik in 1990 (Perona and Malik, 1990). Numerous techniques have been developed for surface fairing and mesh regularization (Sochen et al., 1998; Malladi and Ravve, 2002; Tang et al., 1999; Taubin, 2000) and surface data smoothing (Andrade et al., 2001; Chung and Taylor, 2004; Cachia et al., 2003a,b; Chung et al., 2005; Joshi et al., 2009). Particularly in brain imaging, isotropic heat diffusion on surfaces has been introduced for subsequent statistical analysis involving the random field theory that assumes an isotropic covariance function as a noise model (Andrade et al., 2001; Chung and Taylor, 2004; Cachia et al., 2003a,b). Since then, isotropic diffusion has been mainly used as a standard smoothing technique. Such diffusion approaches mainly use finite element or finite difference schemes that are known to suffer numerical instability if a sufficiently small step size is not chosen in the forward Euler scheme.

Iterated kernel smoothing is also a widely used method in approximately solving diffusion equations on surfaces (Chung et al., 2005; Han et al., 2006). Iterated kernel smoothing is often used in smoothing various anatomical surface data: cortical curvatures (Luders et al., 2006b; Gaser et al., 2006), cortical thickness (Luders et al., 2006a; Bernal-Rusiel et al., 2008), hippocampus surface (Shen et al., 2006; Zhu et al., 2007) and magnetoencephalography (MEG) (Han et al., 2007) and functional-MRI (Hagler Jr. et al., 2006; Jo et al., 2007) on the brain surface. Due to its simplicity, it is probably the most widely used form of surface data smoothing particularly in brain imaging. In iterated kernel smoothing, kernel weights are spatially adapted to follow the shape of the heat kernel in a discrete fashion along a manifold. In the tangent space of the manifold, the heat kernel can be approximated linearly using the Gaussian kernel for small bandwidth. A kernel with large bandwidth is then constructed iteratively applying the kernel with small bandwidth. However, this process compounds the linearization error at each iteration as we demonstrate in this paper.

In this paper, we propose a new kernel regression framework that constructs the heat kernel analytically using the eigenfunctions of the Laplace-Beltrami (LB) operator, avoiding the need for the linear approximation used in Chung et al. (2005) and Han et al. (2006). The proposed method represents isotropic heat diffusion analytically as a series expansion so it avoids the numerical convergence issues associated with solving the diffusion equations numerically (Andrade et al., 2001; Chung and Taylor, 2004; Joshi et al., 2009). Our framework is different from other existing diffusion-based smoothing methods in that it bypasses the various numerical problems such as numerical instability, slow convergence, and accumulated linearization error.

Although there have recently been a few studies that introduce heat kernel in computer vision and machine learning (Belkin et al., 2006), they mainly use heat kernel to compute shape descriptors (Sun et al., 2009; Bronstein and Kokkinos, 2010); or to define a multi-scale metric (de Goes et al., 2008). These studies did not use heat kernel in regressing functional data on manifolds. This is the first study that uses heat kernel in the form of regression for the subsequent statistical analysis. There have been significant developments in kernel methods in machine learning community (Schölkopf and Smola, 2002; Nilsson et al., 2007; Shawe-Taylor and Cristianini, 2004; Steinke and Hein, 2008; Yger and Rakotomamonjy, 2011); however, as far as we know, heat kernel was never used in such frameworks. Furthermore, most kernel methods in machine learning deal with the linear combination of kernels as a solution to penalized regressions, which significantly differs from our kernel regression framework which does not have a penalized cost function.

Recently, wavelets have been popularized for surface and graph data. Spherical wavelets have been used on brain surface data that have been mapped onto a sphere (Nain et al., 2007; Bernal-Rusiel et al., 2008). Since wavelet basis has local supports in both space and scale, the wavelet coefficients from the scale-space decomposition using the spherical wavelets provides shape features that describe local shape variation at a variety of scales and spatial locations. However, spherical wavelets have an intrinsic problem that requires to establish a smooth mapping from the surface to a unit sphere. The spherical mapping introduces a serious metric distortion, which usually compounds subsequent statistical parametric maps (SPM). Furthermore, such basis functions are only orthonormal for data defined on the sphere and result in a less parsimonious representation for data defined on other surfaces compared to the intrinsic LB-eigenfunction expansion (Seo and Chung, 2011). To remedy the limitations of spherical wavelets, the diffusion wavelet transform on graph data structures has been proposed (Antoine et al., 2010; Coifman and Maggioni, 2006; Hammond et al., 2011; Kim et al., 2012).

The primary methodical contribution of this study is the establishment of a unified regression framework that combines the diffusion-, kernel- and wavelet-based methods for scalar data defined on manifolds. Although diffusion-, kernel- and wavelet-based methods seem to be all different methodologies, we can establish a unified framework that relates all of them in a coherent mathematical fashion. This paper extends the conference paper in Kim et al. (2011), where the heat kernel smoothing was introduced to smooth out surface noises in the hippocampus and amygdala. We provide detailed theoretical justification and validation of the proposed unified kernel regression framework. Although the idea of diffusion wavelet transform for surface mesh was explored in Kim et al. (2012), the relationship between the wavelet transform and the proposed kernel regression was not investigated. For the first time, the mathematical equivalence between the two constructs is explained.

The proposed kernel regression framework is subsequently applied in characterizing the growth pattern of mandible surfaces obtained in CT and identifying the regions of mandible that show the most significant localized growth. The length of the displacement vector field is regressed over the mandible surface for increasing the signal to noise ratio and hence statistical sensitivity. To our knowledge, this is the first growth modeling of mandible surface in a continuous fashion without using anatomic landmarks.

2. Methods

2.1. Isotropic Diffusion on Manifolds

Consider a functional measurement $Y(p)$ observed at each point p on a compact manifold $\mathcal{M} \subset \mathbb{R}^3$. We assume the following linear model on Y:

$$Y(p) = \theta(p) + \epsilon(p), \tag{1}$$

where $\theta(p)$ is the unknown mean signal to be estimated and $\epsilon(p)$ is a zero-mean Gaussian random field. We may assume further $Y \in L^2(\mathcal{M})$, the space of square integrable functions on $\mathcal{M}$ with the inner product

$$\langle f, g \rangle = \int_{\mathcal{M}} f(p)g(p)\, d\mu(p), \tag{2}$$

where μ is the Lebesgue measure such that $\mu(\mathcal{M})$ is the total area of $\mathcal{M}$.

Various functional data such as electroencephalography (EEG), magnetoencephalography (MEG) (Han et al., 2007) and functional-MRI (Hagler Jr. et al., 2006; Jo et al., 2007), and anatomical data such as cortical curvatures (Luders et al., 2006b; Gaser et al., 2006), cortical thickness (Luders et al., 2006a; Bernal-Rusiel et al., 2008) and surface coordinates (Chung et al., 2005) can be considered as possible functional measurements. Functional measurements are expected to be noisy and require filtering to boost signal.

Surface measurements have been often filtered using the isotropic diffusion equation of the form (Andrade et al., 2001; Chung, 2001; Cachia et al., 2003a; Rosenberg, 1997):

$$\frac{\partial f}{\partial \sigma} = \Delta f,\ f(p, \sigma = 0) = Y(p), \tag{3}$$

where Δ is the Laplace-Beltrami operator defined on manifold $\mathcal{M}$. The diffusion time σ controls the amount of smoothing. It can be shown that the unique solution of (3) is given by kernel convolution. This can be easily seen as follows.

A Green's function or a fundamental solution of the Cauchy problem (3) is given by the solution of the following equation

$$\frac{\partial f}{\partial \sigma} = \Delta f,\ f(p, \sigma = 0) = \delta(p), \tag{4}$$

where δ is the Dirac delta function. The heat kernel K_σ is a Green's function of (4) (Evans, 1998), i.e.

$$\frac{\partial K_\sigma}{\partial \sigma} = \Delta K_\sigma,\ K_\sigma(p, \sigma = 0) = \delta(p).$$

Since the differential operators are linear in (4), we can further convolve the terms with the initial data Y such that

$$\frac{\partial}{\partial \sigma}(K_\sigma * Y) = \Delta(K_\sigma * Y),\ K_\sigma * Y(p, \sigma = 0) = Y(p),$$

where

$$K_\sigma * Y(p) = \int_{\mathcal{M}} K_\sigma(p, q) Y(q)\, d\mu(q).$$

Hence $K_\sigma * Y$ is a solution of (3).

2.2. Diffusion Smoothing

The isotropic diffusion (3) has been numerically solved by various numerical techniques (Chung, 2001; Andrade et al., 2001; Cachia et al., 2003a,b; Chung and Taylor, 2004). For diffusion smoothing, the diffusion equation needs to be discretized using the cotan formulation (Chung, 2001; Chung and Taylor, 2004; Qiu et al., 2006). Since there are many different cotan formulations, we followed the formulation first given in Chung (2001). Diffusion equation (3) is discretized as

$$\frac{\partial \mathbf{f}}{\partial \sigma} = -\mathbf{A}^{-1}\mathbf{C}\mathbf{f}, \tag{5}$$

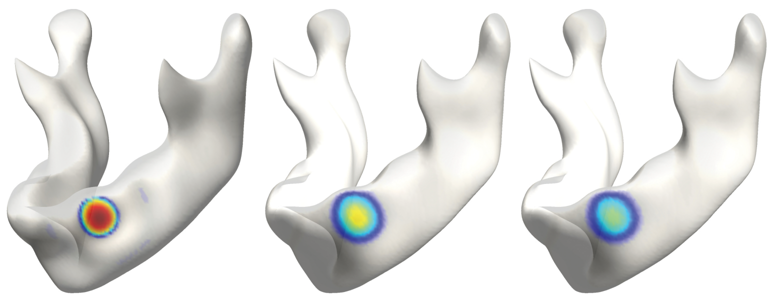

Figure 1: Heat kernel shape with bandwidths 0.025, 1.25 and 5 on a mandible surface. The level sets of the heat kernel form geodesic circles.

where $\mathbf{f} = (f(p_1, \sigma), \cdots, f(p_n, \sigma))'$ is the vector of measurements over all mesh vertices at time σ. $\mathbf{A} = (A_{ij})$ is the stiffness matrix and $\mathbf{C} = (C_{ij})$ is the global coefficient matrix, which is the assemblage of individual element coefficients. The sparse matrices $\mathbf{A}$ and $\mathbf{C}$ are explicitly given as follows.

Let T_{ij}^- and T_{ij}^+ denote two triangles sharing the vertex p_i and its neighboring vertex p_j in a mesh. Let two angles opposite to the edge containing p_i and p_j be ϕ_{ij} and θ_{ij} respectively for T_{ij}^+ and T_{ij}^-. The off-diagonal entries of the stiffness matrix are

$$A_{ij} = \frac{1}{12}\big(|T_{ij}^+| + |T_{ij}^-|\big)$$

if p_i and p_j are adjacent and $A_{ij} = 0$ otherwise. $|\cdot|$ denotes the area of a triangle. The diagonal entries are summed as $A_{ii} = \sum_{j=1}^n A_{ij}$. The off-diagonal entries of the global coefficient matrix are

$$C_{ij} = -\frac{1}{2}(\cot\theta_{ij} + \cot\phi_{ij})$$

if p_i and p_j are adjacent and $C_{ij} = 0$ otherwise. The diagonal entries is similarly given as the sum $C_{ii} = -\sum_{j=1}^n C_{ij}$.

Ordinary differential equation (5) is then further discretized at each point using the forward finite difference scheme: :

$$\mathbf{f}(p_i, \sigma_{n+1}) = \mathbf{f}(p_i, \sigma_n) + (\sigma_{n+1} - \sigma_n)\widehat{\Delta} f(p_i, \sigma_n), \tag{6}$$

where $\widehat{\Delta} f(p_i, \sigma_n)$ is the estimated Laplacian obtained from the i-th row of $-\mathbf{A}^{-1}\mathbf{Cf}$. For the forward Euler scheme to converge, we need to have sufficiently small step size $\Delta\sigma = \sigma_{n+1} - \sigma_n$ for convergence (Chung, 2001).

2.3. Iterated Kernel Smoothing

The diffusion equation (3) can be approximately solved by iteratively performing Gaussian kernel smoothing (Chung et al., 2005). In iterated kernel smoothing, the weights of the kernel are spatially adapted to follow the shape of heat kernel in discrete fashion along a surface mesh. Heat kernel smoothing with large bandwidth can be broken into iterated smoothing with smaller bandwidths (Chung et al., 2005):

$$K_{m\sigma} * Y = \underbrace{K_\sigma * \cdots * K_\sigma}_{m \text{ times}} * Y. \tag{7}$$

Then using the *parametrix expansion* (Rosenberg, 1997; Wang, 1997), we approximate the heat kernel with small bandwidth locally using the Gaussian kernel:

$$K_\sigma(p,q) = \frac{1}{\sqrt{4\pi\sigma}} \exp[-\frac{d^2(p,q)}{4\sigma}][1 + O(\sigma^2)], \tag{8}$$

where $d(p,q)$ is the geodesic distance between p and q. For sufficiently small bandwidth σ, all the kernel weights are concentrated near the center, so the first neighbors of a given mesh vertex is sufficient for approximation. Unfortunately, this approximation is bound to compound error at each additional iteration. For numerical implementation, we used the normalized truncated kernel given by

$$W_\sigma(p,q_i) = \frac{\exp\big[-\frac{d^2(p,q_i)}{4\sigma}\big]}{\sum_{j=0}^{r} \exp\big[-\frac{d^2(p,q_j)}{4\sigma}\big]}, \tag{9}$$

where $q_1, \cdots, q_r$ are r neighboring vertices of $p = q_0$. Denote the truncated kernel convolution as

$$W_\sigma * Y(p) = \sum_{i=0}^{r} W_\sigma(p,q_i) Y(q_i). \tag{10}$$

Then, the iterated heat kernel smoothing is defined as

$$W_{m\sigma} * Y(p) = \underbrace{W_\sigma * \cdots * W_\sigma}_{m \text{ times}} * Y(p).$$

2.4. Heat Kernel Regression

We present a new regression framework for solving the isotropic diffusion equation (3). Let Δ be the Laplace-Beltrami operator on $\mathcal{M}$. Solving the eigenvalue equation

$$\Delta\psi_j = -\lambda\psi_j, \tag{11}$$

we order eigenvalues

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

and corresponding eigenfunctions $\psi_0, \psi_1, \psi_2, \cdots$ (Rosenberg, 1997; Chung et al., 2005; Lévy, 2006; Shi et al., 2009). Then, the eigenfunctions ψ_j form an orthonormal basis in $L^2(\mathcal{M})$. There is extensive literature on the use of eigenvalues and eigenfunctions of the Laplace-Beltrami operator in medical imaging and computer vision (Lévy, 2006; Qiu et al., 2006; Reuter et al., 2009; Reuter, 2010; Zhang et al., 2007, 2010). The eigenvalues have been used in caudate shape discriminators (Niethammer et al., 2007). Qiu et al. used eigenfunctions in constructing splines on cortical surfaces (Qiu et al., 2006). Reuter used the topological features of eigenfunctions (Reuter, 2010). Shi et al. used the Reeb graph of the second eigenfunction in shape characterization and landmark detection in cortical and subcortical structures (Shi et al., 2008, 2009). Lai et al. used the critical points of the second eigenfunction as anatomical landmarks for colon surfaces (Lai et al., 2010). Since the direct application of eigenvalues and eigenfunctions as features of interest is the beyond the scope of the paper, we will not pursue the issue in detail here.

Using the eigenfunctions, *heat kernel* $K_\sigma(p, q)$ is defined as

$$K_\sigma(p, q) = \sum_{j=0}^{\infty} e^{-\lambda_j \sigma} \psi_j(p) \psi_j(q), \tag{12}$$

where σ is the bandwidth of the kernel. Figure 1 shows examples of a heat kernel with different bandwidths. Then *heat kernel regression* or smoothing of functional measurement Y is defined as

$$K_\sigma * Y(p) = \sum_{j=0}^{\infty} e^{-\lambda_j \sigma} \beta_j \psi_j(p), \tag{13}$$

where $\beta_j = \langle Y, \psi_j \rangle$ are Fourier coefficients (Chung et al., 2005) (Figure 2). 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.

Unlike previous approaches to heat diffusion (Andrade et al., 2001; Chung and Taylor, 2004; Joshi et al., 2009; Tasdizen et al., 2006), our proposed method avoids the direct numerical discretization of the diffusion equation. Instead we discretize the basis functions of the given manifold $\mathcal{M}$ by solving for the eigensystem (11) and obtain λ_j and ψ_j.

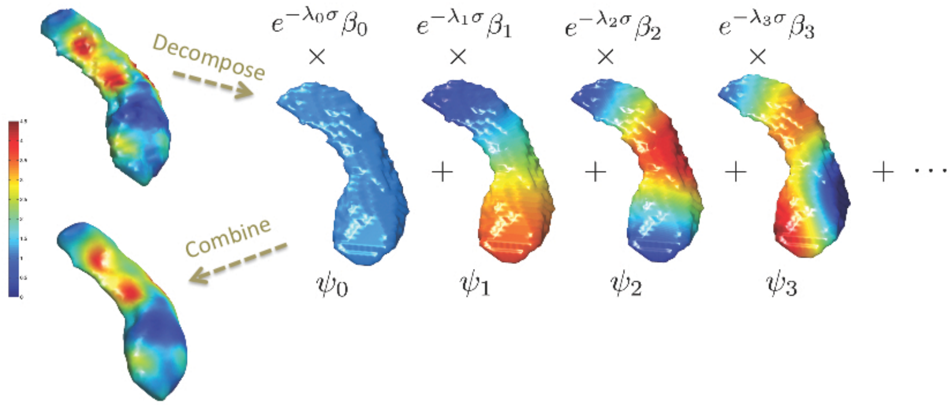

Figure 2: Schematic of heat kernel smoothing. Given a functional data on a surface, we compute the eigenfunctions ψ_j and the Fourier coefficients β_j. Then we combine all the terms and reconstruct the functional signal back.

2.5. Diffusion Wavelet Transform

We can establish the relationship between the kernel regression and a recently popular diffusion wavelet framework. In fact it can be shown that the proposed kernel regression is equivalent to the wavelet transform. This mathematical equivalence removes a need for constructing wavelets using a complicated computational machinery as often done in previous studies (Antoine et al., 2010; Hammond et al., 2011; Kim et al., 2012) and offers a simpler but more unified alternative.

Consider a wavelet basis $W_{\sigma,q}(p)$ obtained from a mother wavelet W with scale and translation parameters σ and q respectively in a Euclidean space:

$$W_{\sigma,q}(p) = \frac{1}{\sigma} W\Big(\frac{p-q}{\sigma}\Big).$$

Generalizing the idea of scaling a mother wavelet in Euclidean space to a curved surface is trivial. However, the difficulty arises when one tries to translate a mother wavelet on a curved surface since it is unclear how to define translation along the surface. If one tries to modify the existing spherical wavelet framework to an arbitrary surface (Nain et al., 2007; Bernal-Rusiel et al., 2008), one immediately encounters the problem of establishing regular grids on an arbitrary surface. The recent works based on the diffusion wavelets bypass this problem by taking bivariate kernel as a mother wavelet (Antoine et al., 2010; Hammond et al., 2011; Mahadevan and Maggioni, 2006; Kim et al., 2012).

For some scale function g that satisfies the admissibility conditions, diffusion wavelet $W_{\sigma,q}(p)$ at position p and scale σ is given by

$$W_{\sigma,q}(p) = \sum_{j=0}^{k} g(\lambda_j \sigma)\psi_j(p)\psi_j(q),$$

where λ_j and ϕ_j are eigenvalues and eigenfunctions of the Laplace-Beltrami operator. The wavelet transform is then given by

$$\langle W_{\sigma,q}, Y\rangle = \int_{\mathcal{M}} W_{\sigma,q}(p)Y(p)\, d\mu(p). \tag{14}$$

If we let $g(\lambda_j\sigma) = \exp(-\lambda_j\sigma)$, we have the heat kernel as the wavelet, i.e.

$$W_{\sigma,p}(q) = H_\sigma(p,q),$$

The bandwidth σ of the heat kernel is the scale parameter while the translation is done by shifting one argument in the bivariate heat kernel. Subsequently, the wavelet transform (14) can be rewritten as

$$\langle W_{\sigma,p}, Y\rangle = \sum_{j=0}^{k} e^{-\lambda_j\sigma}\beta_j\psi_j(q) \tag{15}$$

with $\beta_j = \langle Y, \psi_j\rangle$. The expression (15) is exactly the finite truncation of heat kernel regression in (12). Hence, diffusion wavelet analysis can be simply performed within the proposed heat kernel regression framework without any additional wavelet machinery. From now on, we will not distinguish heat kernel regression and diffusion wavelet transform.

Although the heat kernel regression is constructed using global basis functions ψ_j, surprisingly the kernel regression at each point p coincides with the wavelet transform at that point. Hence, it also inherits all the localization property of wavelets at that point. This is clearly demonstrated in a simulation given in Figure 3, where a step function of value 1 in the circular band $1/8 < \theta < 1/4$ (angle from the north pole) and of value 0 outside of the band is constructed. Note that on a sphere, the Laplace-Beltrami operator is the spherical Laplacian and its eigenfunctions are spherical harmonics Y_{lm} of degree l and order m. Then the step function is reconstructed using the spherical harmonic series expansion

$$Y(p) = \sum_{l=0}^{78}\sum_{m=-l}^{l} \beta_{lm}Y_{lm}(p),$$

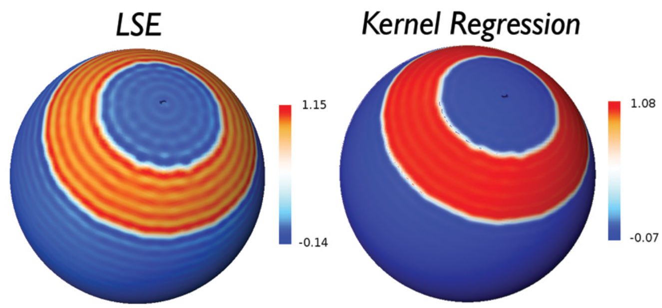

Figure 3: Gibbs phenomenon (ringing artifacts) is visible in the spherical harmonic series expansion with degree 78 via LSE of the step function defined on a sphere. On the other hand, the heat kernel regression with the same degree and bandwidth 0.0001 shows less visible artifacts.

where the spherical harmonic coefficients $\beta_{lm} = \langle Y, Y_{lm} \rangle$ are obtained by the least squares estimation (LSE). On the unit sphere, we used the heat kernel regression of the form

$$Y(p) = \sum_{l=0}^{78} \sum_{m=-l}^{l} e^{-l(l+1)\sigma} \beta_{lm} Y_{lm}(p)$$

with the small bandwidth $\sigma = 0.0001$. The spherical harmonic expansion clearly shows severe ringing artifacts compared to the kernel regression, which inherits the localization power of wavelets. This is why the Gibbs phenomenon is not visibly significant.

2.6. Parameter Estimation in Heat Kernel Smoothing

Since the closed form expression for the eigenfunctions of the Laplace-Beltrami operator on an arbitrary surface is unknown, the eigenfunctions are numerically computed by discretizing the Laplace-Beltrami operator. To solve the eigensystem (11), we need to discretize it on mandible triangular meshes using the cotan formulation (Chung, 2001; Chung and Taylor, 2004; Shi et al., 2009; Qiu et al., 2006; Lévy, 2006; Reuter et al., 2006, 2009; Rustamov, 2007; Zhang et al., 2007; Vallet and Lévy, 2008; Wardetzky, 2008).

Among many different cotan formulations used in computer vision and medical image analysis, we used the formulation given in Chung (2001) and Qiu et al. (2006). It requires discretizing (11) as the following generalized eigenvalue problem:

$$\mathbf{C}\psi = \lambda \mathbf{A}\psi, \tag{16}$$

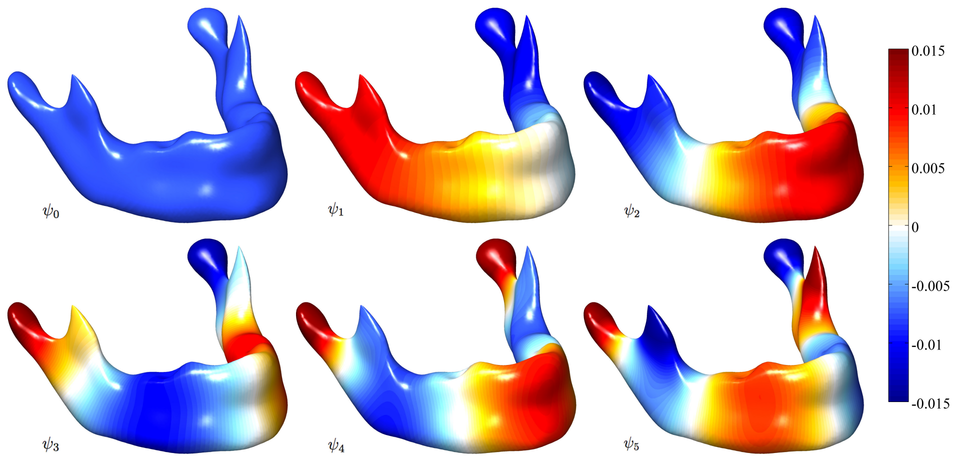

Figure 4: Eigenfunctions of various degrees for a sample mandible surface. The eigenfunctions are projected on the surface smoothed by the proposed heat kernel smoothing with bandwidth $\sigma = 0.5$ and degree $k = 132$. The smoothed surface is obtained by heat kernel smoothing applied to the coordinates of the surface mesh with the same parameter while preserving the topology of mesh. The first eigenfunction is simply $\psi_0 = 1/\sqrt{\mu(\mathcal{M})}$. The color scale is thresholded at ± 0.015 for better visualization.

where the global coefficient matrix $\mathbf{C}$ is the assemblage of individual element coefficients and $\mathcal{A}$ are the stiffness matrix. We solved (16) using the *Implicitly Restarted Arnoldi Method* (Hernandez et al., 2006; Lehoucq et al., 1998) without consuming large amount of memory and time for sparse entries. Figure 4 shows the first few eigenfunctions for a mandible surface. The first eigenfunction is trivially given as $\lambda_0 = 0$ and $\psi_0 = 1/\sqrt{\mu(\mathcal{M})}$ for a closed compact surface. It is possible to have multiple eigenfunctions corresponding to a single eigenvalue.

Once we obtain the eigenfunctions numerically, we estimate the kernel regression parameters β_j by minimizing the sum of squared residual using the least squares estimation (LSE):

$$\arg \min_{\beta_0, \cdots, \beta_k} \left\| Y(p) - \sum_{j=0}^{k} e^{-\lambda_j \sigma} \beta_j \psi_j(p) \right\|^2. \tag{17}$$

The least squares method is often used in estimating the coefficients in spherical harmonic expansion (Shen et al., 2004; Styner et al., 2006; Chung et al., 2008). Suppose we have n mesh vertices $p_1, \cdots, p_n$. Let

$$\mathbf{Y} = (Y(p_1), \cdots, Y(p_n))'$$

be the surface measurements over all n vertices. Denote the j-th eigenfunction evaluated at n vertices as

$$\mathbf{\Psi}_j = (\psi_j(p_1), \cdots, \psi_j(p_n))'.$$

By letting $\sigma = 0$, (17) achieves the minimum when

$$\mathbf{Y} = \mathbf{\Psi}\boldsymbol{\beta}, \tag{18}$$

where $\mathbf{\Psi} = (\mathbf{\Psi}_0, \cdots, \mathbf{\Psi}_k)$ is the matrix of size $n \times (k+1)$. The LSE estimation of coefficients $\boldsymbol{\beta}$ is then given by

$$\widehat{\boldsymbol{\beta}} = (\mathbf{\Psi}'\mathbf{\Psi})^{-1}\mathbf{\Psi}'\mathbf{Y}. \tag{19}$$

Since it is expected that the number of mesh vertices is substantially lager than the number of eigenfunctions to be used, $\mathbf{\Psi}'\mathbf{\Psi}$ is well conditioned and invertible.

2.7. Random Field Theory

Once we smooth functional data on a surface, we apply the statistical parametric mapping (SPM) framework for analyzing and visualizing statistical tests performed on the template surface that is often used in structural neuroimaging studies (Andrade et al., 2001; Lerch and Evans, 2005; Wang et al., 2010; Worlsey et al., 1995; Yushkevich et al., 2008). Since test statistics are constructed over all mesh vertices on the mandible, multiple comparisons need to be accounted for possibly using the random field theory (Taylor and Worsley, 2007; Worlsey et al., 1995; Worsley et al., 2004). The random field theory assumes the measurements to be smooth Gaussian random field. Heat kernel smoothing will make data more smooth and Gaussian and increase the signal-to-noise ratio (Chung et al., 2005). The proposed kernel smoothing can be naturally integrated into the random field theory based statistical inference framework (Taylor and Worsley, 2007; Worsley et al., 2004; Worlsey et al., 1995).

Given linear model (1), we are interested in determining the significance of θ, i.e.

$$H_0 : \theta(p) = 0 \text{ for all } p \in \mathcal{M} \quad \text{vs.} \quad H_1 : \theta(p) > 0 \text{ for some } x \in \mathcal{M}. \tag{20}$$

Note that any point p_0 that gives $\theta(p_0) > 0$ is considered as signal. The hypotheses (20) are an infinite dimensional multiple comparisons problem for continuously indexed hypotheses over the manifold $\mathcal{M}$. The underlying group level signal h is estimated using the proposed heat kernel regression.

Subsequently, a test statistic is often given by a T- or F-field $Y(p)$ (Worsley et al., 2004; Worlsey et al., 1995).

The multiple comparison corrected p-value computation is then given by the random field theory (Adler, 1981; Cao and Worsley, 2001; Taylor and Worsley, 2007; Worsley, 2003). For the F-field Y with α and β degrees of freedom defined on 2D manifolds $\mathcal{M}_F$, it is known that

$$P\Big(\sup_{p\in\mathcal{M}_F} Y(p) > h\Big) \approx \mu_2(\mathcal{M}_T)\rho_2(h) + \mu_0(\mathcal{M}_F)\rho_0(h) \tag{21}$$

for sufficiently high threshold h. $\mu_d(\mathcal{M}_F)$ is the d-th Minkowski functional of $\mathcal{M}_F$ and ρ_d is the d-th Euler characteristic (EC) density of Y. The Minkowski functionals are simply

$$\begin{aligned} \mu_2(\mathcal{M}_T) &= \text{area}(\mathcal{M}_T)/2 \\ \mu_0(\mathcal{M}_T) &= \chi(\mathcal{M}_T) = 2. \end{aligned}$$

The EC-density for F-field is then given by

$$\begin{aligned} \rho_2 &= \frac{1}{4\pi\sigma^2}\frac{\Gamma(\frac{\alpha+\beta-2}{2})}{\Gamma(\frac{\alpha}{2})\Gamma(\frac{\beta}{2})}\Big(\frac{\alpha h}{\beta}\Big)^{\frac{(\alpha-2)}{2}}\Big(1+\frac{\alpha h}{\beta}\Big)^{-\frac{(\alpha+\beta-2)}{2}}\Big[(\beta-1)\frac{\alpha h}{\beta}-(\alpha-1)\Big] \\ \rho_0 &= 1 - P(F_{\alpha,\beta} \le h), \end{aligned}$$

where $P(F_{\alpha,\beta} \le h)$ is the cumulative distribution function of F-stat with α and β degrees of freedom. Note that the second order term $\mu_2(\mathcal{M}_T)\rho_2(h)$ dominates the expression (21) and it explicitly has the bandwidth σ of the kernel regression.

3. Experiments

3.1. CT Image Preprocessing

We applied the proposed smoothing method to mandible surfaces obtained from CT. The CT images were obtained from several different models of GE multi-slice helical CT scanners. The CT scans were acquired directly in the axial plane with 1.25 mm slice thickness, matrix size of 512×512 and 15–25 cm field of view (FOV). Image resolution varied as voxel size ranged from 0.25 mm^3 to 0.49 mm^3 as determined by the ratio of FOV divided by the matrix. CT scans were converted to DICOM format and subsequently Analyze 8.1 software package (AnalyzeDirect, Inc., Overland Park, KS) was used in segmenting binary mandible structure based on histogram thresholding.

Image acquisition and processing artifacts, and partial voluming produce topological defects such as holes and handles in any medical images. In mandibles CT images, unwanted cavities, holes and handles in the binary segmentation are mainly resulting from differences in CT intensity between relatively low mandible and teeth and more dense cortical bone and the interior trabecular bone (Andresen et al., 2000; Loubele et al., 2006). In mandibles, these topological noises can appear in thin or cancellous bone, such as in the condylar head and posterior palate (Stratemann et al., 2010). An example is shown in Figure 5 where the tooth cavity forms a bridge over the mandible. If we apply the isosurface extraction on the topologically defect segmentation results, the resulting surface will have many tiny handles (Wood et al., 2004; Yotter et al., 2009). These handles complicate subsequent surface mesh operations such as smoothing and parameterization. So it is necessary to correct the topology by filling the holes and removing handles. If we correct such topological defects, it is expected that the resulting isosurface is topologically equivalent to a sphere.

Various topology correction techniques have been proposed in medical image processing. Rather than attempting to repair the topological defects of the already extracted surfaces (Wood et al., 2004; Yotter et al., 2009), we performed the topological simplification on the volume representation directly using morphological operations (Guskov and Wood, 2001; Van Den Boomgaard and Van Balen, 1992; Yotter et al., 2009). The direct correction on surface meshes can possibly cause surfaces to intersect each other (Wood et al., 2004). By checking the Euler characteristic, the holes were automatically filled up using morphological operations to make the mandible binary volume to be topologically equivalent to a solid sphere. All areas enclosed by the higher density bone that is included in the mandible definition are morphed into being included as part of the definition of the mandible object. Then, the hole-filled images were converted to surface meshes via the marching cubes algorithm.

In our fully automated algorithm, we first removed the speckles of noise components by identifying the largest connected component in the binary volume. Then we applied the morphological closing operation in each 2D slice of CT images one by one in all three axes. Recombining the topology corrected 2D slices resulted in topologically correct surface meshes (Figure 5). The reason 2D topological closing operations are used is mainly due to its performance and relatively simpler implementation than 3D topological closing operations. In 2D topological operations, we only need to worry about 8 neighboring voxels in a 2D image slice. In contrast, we need to worry about 26 neighboring voxels in a 3D image volume. There are many

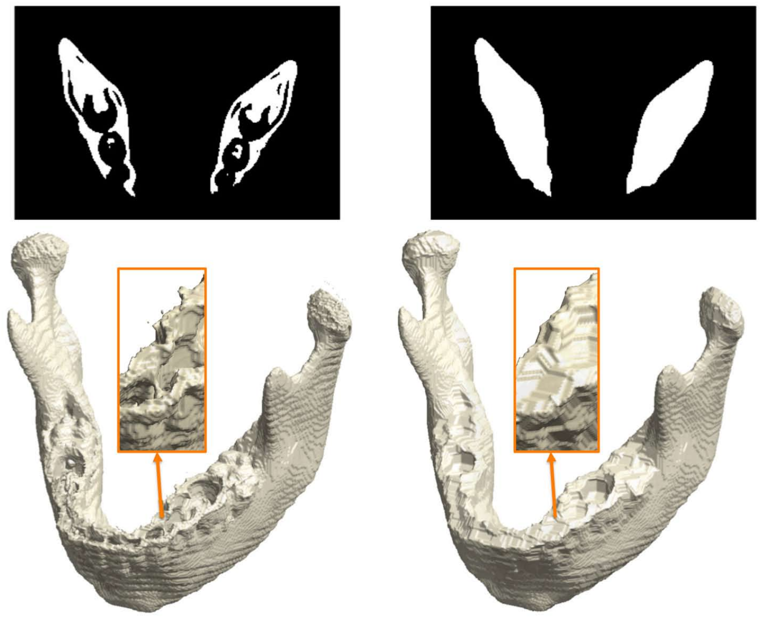

Figure 5: Topological correction on mandible binary segmentation and surface. Disjoint tiny speckles of noisy components are removed by labeling the largest connected component, and holes and handles are removed by the morphological closing operation. Left: A slice shows holes and handles in teeth regions. The isosurface has Euler characteristic $\chi = 50$. Right: After the correction with $\chi = 2$.

large concave regions left out by teeth and fillings. These regions may not be closed with 3D closing operations. However, we found that they can be easily patched up with 2D closing operations, which tend to put more constraints on the underlying topology. Instead of performing a single 3D closing operation that may not work, we are sequentially performing 2D closing operations in each image slices in all x-, y- and z-directions. Figure 6 shows a simulated cavity example where existing 3D closing operation will not patch (Van Den Boomgaard and Van Balen, 1992) while the proposed sequential application of 2D closing operations will easily patch. Note that any 3D object, whose every 2D cross-sections are topologically equivalent to a solid disk, is topologically equivalent to a solid sphere. So the problem of 3D topology correction can be reduced to a much simpler problem of 2D topology correction of multiple slices.

At the end of the processing, we checked the Euler characteristic of resulting surface meshes. Note that for each triangle, there are three edges.

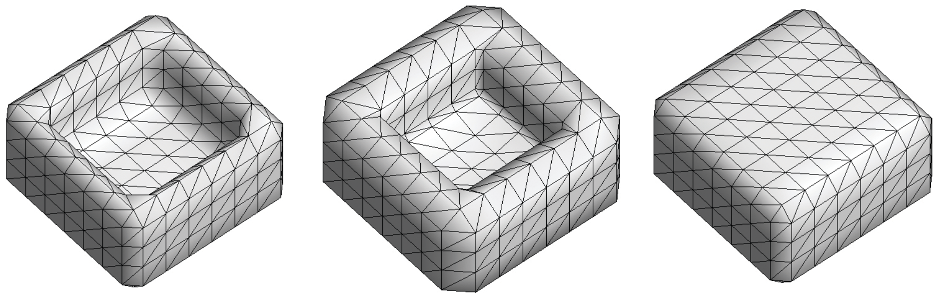

Figure 6: Cavity patching by topological closing operations. Left: Surface model of the binary volume that simulates a tooth cavity. Middle: 3D image volume based closing operation that will not properly patch the cavity region. Right: 2D image slice based closing operation that patch the cavity region properly.

For a closed surface topologically equivalent to a sphere, two adjacent triangles share the same edge. Hence, the total number of edges is $E = 3F/2$. Hence, we checked if the Euler characteristic is simply given by $\chi = V - F/2$ at the end. All binary volumes produced the topologically correct surfaces without an exception. Figure 5 shows an example of before and after the topology correction.

3.2. Validation and Performance Analysis of Heat Kernel Smoothing

Here, we compared the performance of the proposed kernel regression against iterated kernel and diffusion smoothing techniques. The high accuracy of the heat kernel construction using LB-eigenfunctions is previously reported in Kim et al. (2011).

3.2.1. Comparison against iterated kernel smoothing

The proposed heat kernel regression was compared against the widely used iterated kernel smoothing framework (Chung et al., 2005; Hagler Jr. et al., 2006; Han et al., 2006). We compared the performance of iterated kernel smoothing (10) against heat kernel smoothing. The x, y and z surface coordinates are treated as functional measurements on the original surface and smoothed with the both methods. For the comparison of performance between both smoothing methods, we calculated the root mean squared errors (RMSE) between them. The mean of the squared errors is taken over the surface. For the heat kernel regression, we used the bandwidth $\sigma = 0.5$ and eigenfunctions up to $k = 132$ degree. For iterated kernel smoothing,

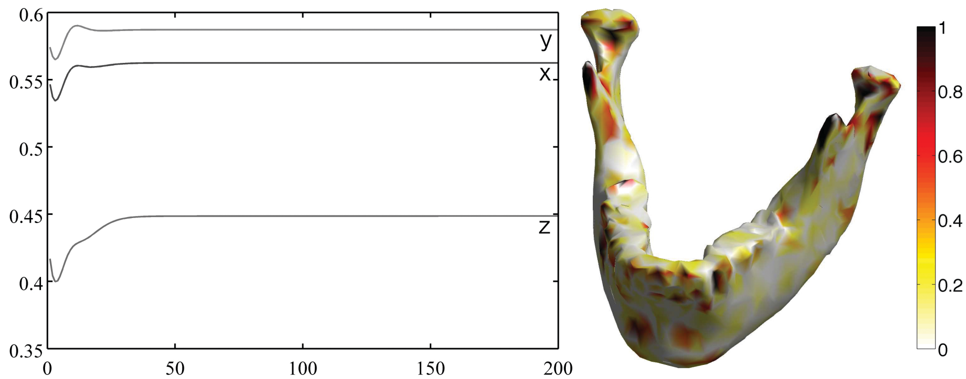

Figure 7: Plot of the RMSE of iterated kernel smoothing against the proposed heat kernel regression for coordinates x (middle), y (top) and z (bottom) over the number of iterations up to 200. For heat kernel regression, $\sigma = 0.5$ and $k = 132$ are used. Widely used iterated kernel smoothing does not converge to heat diffusion but something else. The right figure is the squared difference between the two methods. The difference is mainly localized in high curvature areas, where the Gaussian kernel used in the iterated kernel smoothing fails to approximate the heat kernel.

we varied the number of iterations $1 \leq m \leq 200$ with the correspondingly smaller bandwidth $0.5/m$ to have the effective bandwidth of 0.5. The performance of the iterated kernel smoothing depended on the number of iterations, as shown in the plot of RMSE of mesh coordinates over the number of iterations (Figure 7). The RMSE was up to 0.5901 and it did not decrease even when we increased the number of iterations. The right image in Figure 7 is the squared difference between the two methods. The difference is mainly localized in high curvature areas, where the Gaussian kernel used in the iterated kernel smoothing fails to approximate the heat kernel. This comparison quantitatively demonstrates the limitation of iterated heat kernel smoothing which does not converge to heat diffusion.

3.2.2. Comparison against diffusion smoothing

We further compared the proposed heat kernel regression to diffusion smoothing widely used in smoothing surface data (Andrade et al., 2001; Cachia et al., 2003a,b; Chung and Taylor, 2004). For the forward Euler scheme (6) to converge, we need to have sufficiently small step size $\Delta\sigma$ for convergence. We investigated the convergence of diffusion smoothing against the heat kernel regression with bandwidths $\sigma = 0.5, 20, 50, 100$ and $k = 132$. For diffusion smoothing, small fixed step size of $\Delta\sigma = 0.025$ was used with $m = 20, 800, 2000, 4000$ iterations. The diffusion smoothing result was found

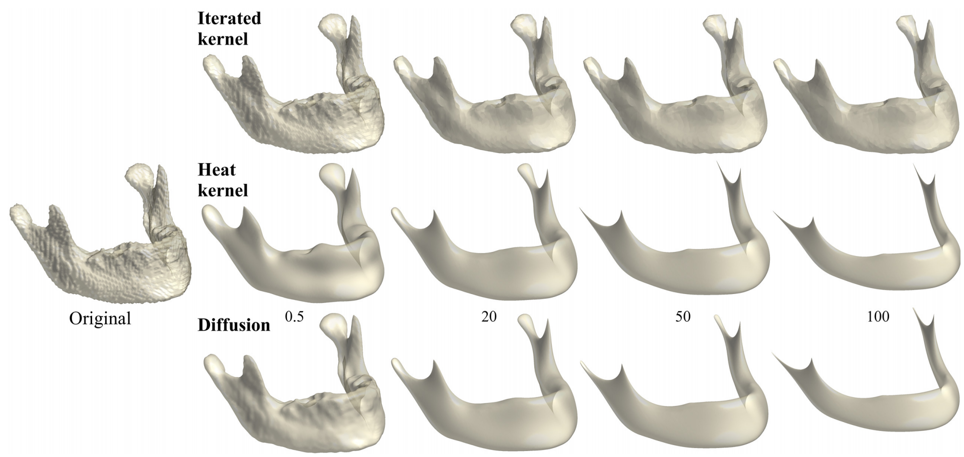

Figure 8: Smoothed mandible surfaces using three different techniques with the same bandwidths. They are all expected to be the solution of isotropic diffusion. The x, y and z surface coordinates are treated as functional measurements on the original surface and smoothed. The proposed heat kernel smoothing is done with various bandwidths, $\sigma = 0.5, 20, 50, 100$. Iterated kernel smoothing performs iterative kernel smoothing with heat kernel approximated linearly with Gaussian kernel (Chung et al., 2005). Diffusion smoothing directly solves the diffusion equation using the same FEM discretization (Chung and Taylor, 2004). The diffusion smoothing and heat kernel smoothing are supposed to converge as the bandwidth increases.

to be inaccurate for less than 10 iterations, but it converged quickly to heat kernel smoothing as the number of iterations m increases and giving the compatible results.

Figure 8 shows the result of smoothing surface coordinates with three different techniques: iterated kernel smoothing (Chung et al., 2005), the proposed kernel regression and diffusion smoothing based on FEM discretization (Chung and Taylor, 2004). We replaced the original surface coordinates with the smoothed ones for the visualization at the end. However, in the actual computation, we did not replace the original surface coordinates for all the three methods. Iterated kernel smoothing compounds the discretization errors over iterations so it does not converge to the kernel regression and diffusion smoothing. Diffusion smoothing and heat kernel smoothing share the same FEM discretization and expected to converge as the bandwidth increases.

3.3. Simulation Studies

Since there is no known ground truth in the imaging data set we are using, it is uncertain how the proposed method will perform in the real data.

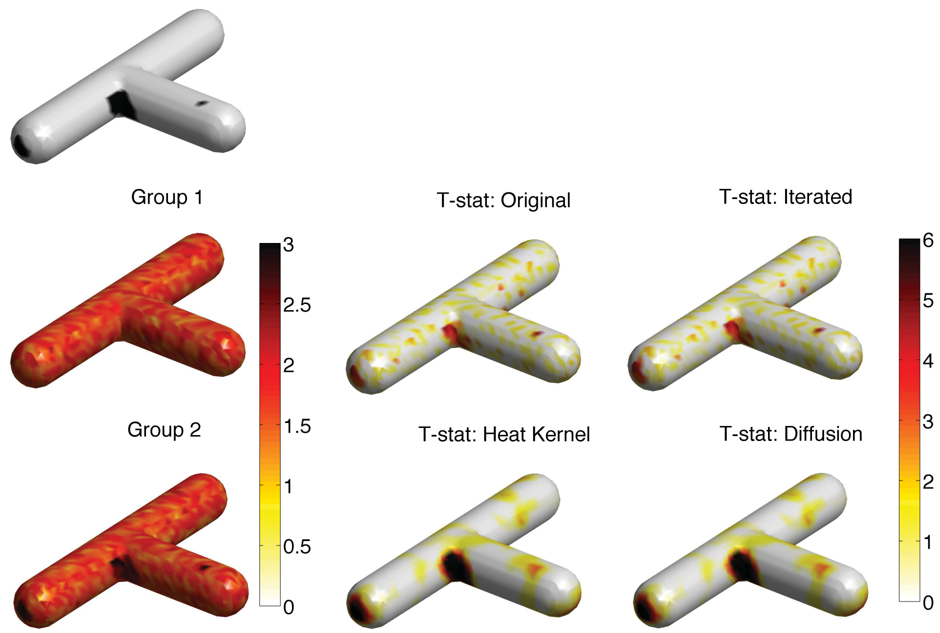

Figure 9: Simulation study I on a T-junction shaped surface where three black signal regions of different sizes are taken as the ground truth. 60 independent functional measurements on the T-junction were simulated as $|N(0, 2^2)|$ at each mesh vertices. We are only simulating positive numbers to better reflect the positive measurements used in the study. Value 1 was added to the black regions in 30 of measurements which served as group 2 while the other 30 measurements were taken as group 1. T-statistics are shown for these simulation (original) and three techniques with bandwidth 0.5. Heat kernel smoothing performed the best in detecting the ground truth.

Therefore, it is necessary to perform simulation studies with the ground truths. We performed two simulations with small and larg signal-to-noise ratio (SNR) settings on a T-junction shaped surface (Figure 9). The T-junction surface is chosen since it is a surface with three different curvatures: convex, concave and almost flat regions. Note surface smoothing methods perform differently under different curvatures. Three black signal regions of different sizes were taken as the ground truth at these regions. 60 independent functional measurements on the T-junction were simulated as $|N(0, \gamma^2)|$, the absolute value of normal distribution with mean 0 and variance γ^2, at each mesh vertices. Subsequently, value 1 was added to the black regions in 30 of measurements which served as group 2 while the other 30 measurements were taken as group 1. So the group 1 has distribution $|N(0, \gamma^2)|$ while the group 2 has distribution $|N(1, \gamma^2)|$ in the signal regions.

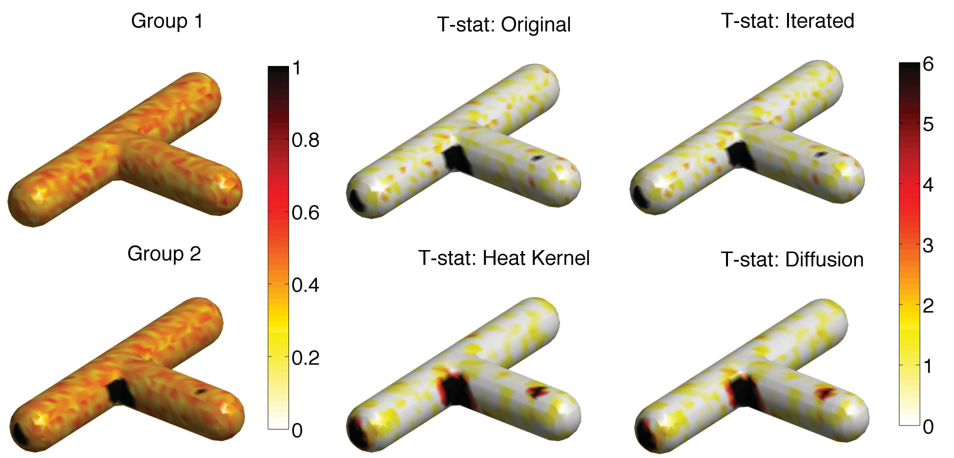

Figure 10: Simulation study II on a T-junction shaped surface where the ground truth is the same as the simulation study I (Figure 9). 60 independent functional measurements on the T-junction were simulated as $|N(0, 0.5^2)|$ at each mesh vertices. Value 1 was added to the black regions in 30 of measurements which served as group 2 while the remaining 30 measurements are taken as group 1. Due to large SNR, the group means show visible group separations. All the methods detected the signal regions; however, the heat kernel smoothing and diffusion smoothing techniques were more sensitive at large SNR.

Larger variance γ^2 corresponds to smaller SNR.

In Study I, $\gamma^2 = 2^2$ was used to simulate functional measurements with substantially smaller SNR. Figure 9 shows the simulation results. For iterated kernel and diffusion smoothing, we used the bandwidth $\sigma = 0.5$ and 100 iterations. For smaller SNR, it is necessary to smooth with larger bandwidth, which is determined empirically. For heat kernel smoothing, the same bandwidth and 1000 eigenfunctions were used. The same number of eigenfunctions is used through the study. For all three smoothing techniques, bandwidth is the main parameter that determines the performance. The results are stable under the perturbation of other parameters. Then we performed the two sample t-test with the random field theory based threshold of 4.90 to detect the group difference at 0.05 level.

Iterated kernel smoothing as well as using no smoothing at all did not correctly identify any signal region. However, heat kernel and diffusion smoothing correctly identified 94% and 91% of the signal regions. This improvement for heat kernel smoothing is significant considering that the error rate of 5% (0.05) is considered as the standard threshold for accepting or rejecting in a hypothesis. Also heat kernel and diffusion smoothing incorrectly identified 0.26% and 0.26% non-signal regions as signal. The both methods use the similar FEM discretization schemes although diffusion smoothing surffer more discretization error. The discretization error is related to the forward Euler scheme that is often employed in diffusion smoothing (Chung, 2001; Andrade et al., 2001; Cachia et al., 2003b). For better accuracy, extremely small time step is required but this requires a very large number of iterations, which slows the method drastically. There might be a more accurate faster discretization scheme but the proposed method is compared against existing standard methods in medical imaging literature that are actually used in practice.

In Study II, $\gamma^2 = 0.5^2$ was used to simulate functional measurements with substantially larger SNR. Due to large SNR, the group means showed visible group separations (Figure 10). For iterated kernel and diffusion smoothing, we used bandwidth $\sigma = 0.1$ and 100 iterations. For heat kernel smoothing, the same bandwidth and with 1000 eigenfunctions were used. All the methods detected the signal regions; however, the heat kernel smoothing and diffusion smoothing techniques are more sensitive at large SNR. All the methods correctly identified the signal regions with 100% accuracy. Iterated kernel smoothing as well as without any smoothing did not incorrectly identified any non-signal regions as signal. However, due to the blurring effects, heat kernel and diffusion smoothing incorrectly identified 0.9% and 0.8% non-signal regions as signal, which is negligible. For large SNR setting, all the methods were reasonably able to detect the correct signal regions without much error.

In summary, in larger SNR, all three methods performed well. However, in substantially smaller SNR, the proposed kernel regression performed the best closely followed by diffusion smoothing. The iterated kernel smoothing as well as no smoothing at all did not perform well in the low SNR setting.

4. Application: Mandible Growth Analysis

As an illustration of the proposed kernel regression technique, we performed a mandible growth analysis on the CT imaging data set consisting of 77 human subjects between ages 0 and 19. Subjects are binned into three age categories: ages between 0 and 6 years (group I), between 7 and 12 years (group II), and between 13 and 19 years (group III). There are 26, 20 and 31 subjects in group I, II and III respectively. The main biological hypothesis of interest is if there are any localized growth spurts between these age groups. Mandible surface meshes for all subjects were constructed through the image acquisition and processing steps explained in the previous section. For surface alignment, diffeomorphic surface registration has been

Figure 11: Left: Mandible F155-12-08 which forms an initial template $\mathcal{M}_I$. All other mandibles are affine registered to F155-12-08. Middle: The superimposition of affine registered mandibles showing local misalignment. Diffeomorphic registration is then performed to warp misaligned affine transformed mandibles. Right: The average of deformation with respect to F155-12-08 provides the final population average template $\mathcal{M}_F$ where statistical parametric maps will be constructed.

performed to align mandible surfaces across subjects (Miller and Qiu, 2009; Vaillant et al., 2007; Qiu and Miller, 2008; Yang et al., 2011).

4.1. Diffeomorphic Surface Registration

We have chosen a 12 year old subject identified as F155-12-08, which served as the reference template in previous studies (Seo et al., 2010, 2011), as initial template $\mathcal{M}_I$ and aligned the remaining 76 mandibles to the initial template affinely to remove the overall size variability. Some subject may have larger mandible than others so it is necessary to remove the global size differences in localized shape modeling. From the affine transformed individual mandible surfaces $\mathcal{M}_j$, we performed an additional nonlinear surface registration to the template using the large deformation diffeomorphic metric mapping (LDDMM) framework (Miller and Qiu, 2009; Vaillant et al., 2007; Qiu and Miller, 2008; Yang et al., 2011).

In LDDMM framework (Miller and Qiu, 2009; Vaillant et al., 2007; Qiu and Miller, 2008; Yang et al., 2011), given a surface $\mathcal{M}$, the metric space is constructed as an orbit of $\mathcal{M}$ under the group of diffeomorphic transformations $\mathcal{G}$. The diffeomorphic transformations (one-to-one, smooth forward and inverse transformation) are introduced as transformations of the coordinates on the background space $\Omega \subset \mathbb{R}^3$, : $\Omega \to \Omega$. The diffeomorphisms $\phi_t \in \mathcal{G}$ is constructed as a flow of ordinary differential equations (ODE), where $\phi_t, t \in [0,1]$ follows

$$\dot{\phi_t} = v_t(\phi_t), \quad \phi_0 = \mathtt{Id}, \quad t \in [0,1], \tag{22}$$

where $\mathtt{Id}$ denotes the identity map and v_t are the associated velocity vector fields. The vector fields v_t are constrained to be sufficiently smooth, so that (22) is integrable and generates diffeomorphic transformations over finite time. The smoothness is ensured by forcing v_t to lie in a smooth vector field V, which is modeled as a reproducing kernel Hilbert space with linear operator L associated with norm $\|u\|_V^2 = \langle Lu, u\rangle_2$ (Dupuis et al., 1998). The group of diffeomorphisms $\mathcal{G}(V)$ is then the solutions of (22) with the vector fields satisfying $\int_0^1 \|v_t\|_V dt < \infty$.

Now, given the template surface $\mathcal{M}$ and an individual surface $\mathcal{M}_j$, the geodesic $\phi_t, t \in [0,1]$ which lies in the manifold of diffeomorphisms and connects $\mathcal{M}$ and $\mathcal{M}_j$, is defined as

$$\phi_0 = \mathtt{Id}, \quad \phi_1 \cdot \mathcal{M} = \mathcal{M}_j.$$

For our application, we employed the LDDMM approach to estimate the template among all subjects. The estimated template can be simply computed through averaging the initial velocity across all the subjects (Zhong and Qiu, 2010), which is similar to the unbiased template estimation approach in Joshi et al. (2004). Subsequently, we recomputed the displacement fields with respect to the initial template $\mathcal{M}_I$, which is a 12 year old subject identified as F155-12-08. We then averaged the deformation fields from the initial template $\mathcal{M}_I$ to individual subjects, we obtain the yet another final template $\mathcal{M}_F$. Figure 11 shows the initial and final templates. Figure 12 shows the mean displacement differences between the groups I and II (top) and II and III (bottom). Each row shows the group differences of the displacement: group II - group I (first row) and group III - group II (second row). The arrows are the growth direction given by the mean displacement differences and colors indicate their lengths in mm.

4.2. Statistical Analysis

We are interested in determining the significance of the mean displacement differences in Figure 12. Since the length measurement provides a much easier biological interpretation, we used the length of displacement vector as a response variable. The random field theory assumes the measurements to be smooth Gaussian random field. Heat kernel smoothing on the length measurement will make the length measurement more smooth and Gaussian as well as increase the signal-to-noise ratio (Chung et al., 2005). Heat kernel smoothing is applied with bandwidth $\sigma = 20$ using 1000 eigenfunctions on the final template $\mathcal{M}_F$. The number of eigenfunctions used is more than sufficient to guarantee relative error less than 0.3%. The

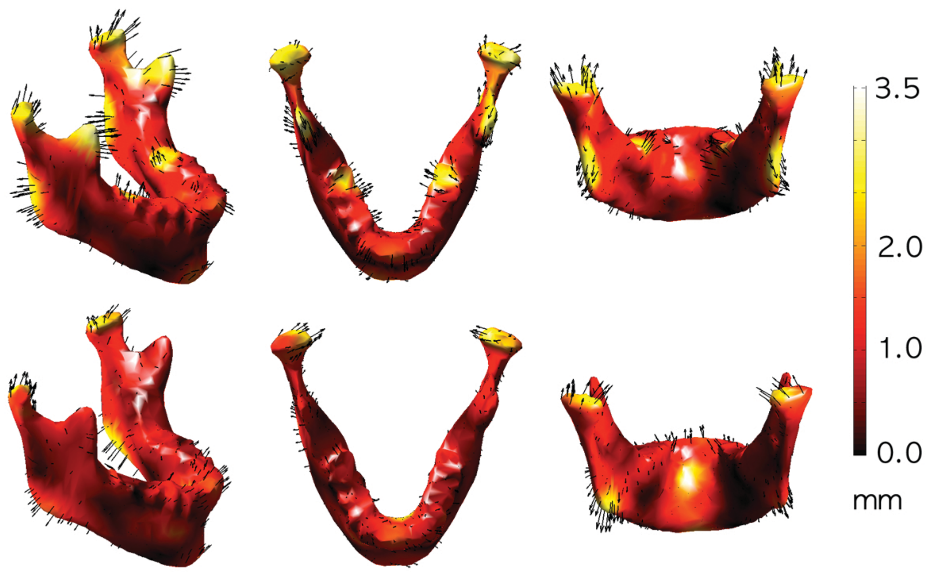

Figure 12: Mandibles are binned into three age groups: group I (ages 0 and 6), group II (ages 7 and 12) and group III (ages 13 and 19). Each row shows the mean group differences of the displacement: group II - group I (first row) and group III - group II (second row). The arrows are the mean displacement differences and colors indicate their lengths in mm.

heat kernel smoothing of the displacement length is taken as the response variable. We constructed the F random field testing the length difference between the age groups I and II, and II and III showing the regions of growth spurts (Figure 13).

For comparing the groups I and II, it is based on F-field with 1 and 44 degrees of freedom while for the groups II and III, it is based on F-field with 1 and 49. The multiple comparison corrected F-stat thresholds corresponding to $\alpha = 0.05$ and 0.01 levels are respectively 8.00 and 10.52 (group II-I) and 8.00 and 10.67 (group III- II). In the F-statistic map shown in Figure 13, any black and red regions are considered as exhibiting growth spurts at 0.01 and 0.05 levels respectively. Our finding is consistent with previous findings of simultaneous growth forward and downward directions (Scott, 1976; Smartt Jr. et al., 2005; Walker and Kowalski, 1972; Lewis et al., 1982; Seo et al., 2011) and also bilateral growth (Enlow and Hans, 1996).

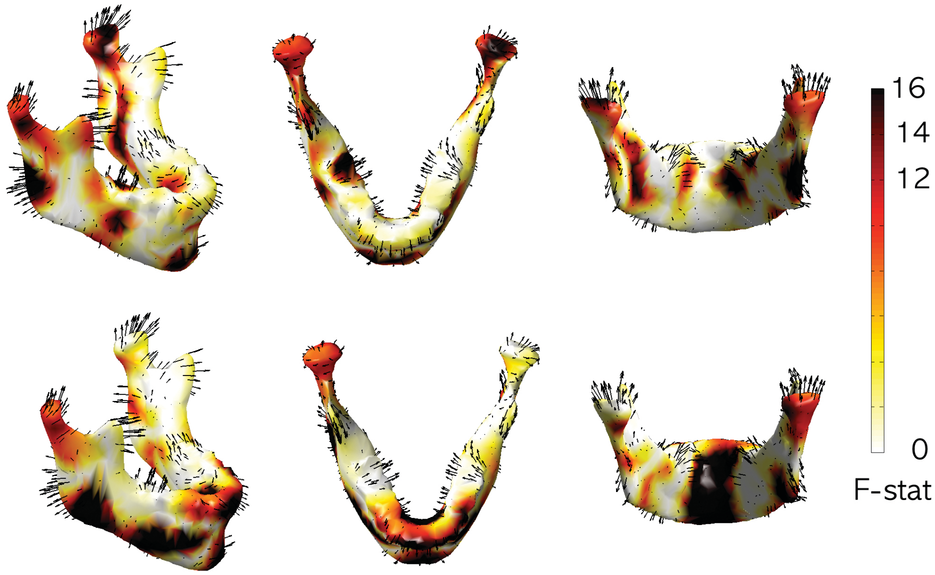

Figure 13: F-statistic map showing the regions of significant mean displacement difference shown in Figure 12.

5. Conclusions

This study presents a novel heat kernel regression framework where the functional measurements are expanded using the weighted Laplace-Beltrami eigenfunctions analytically. The weighted eigenfunction expansion is related to isotropic heat diffusion and the diffusion wavelet transform. The method is validated on a unit sphere, where the spherical harmonics are the eigenfunctions. As demonstrated in the validation, the heat kernel regression provides more accurate results in comparison with the other surface-based smoothing techniques. However, in terms of statistical performance, as shown in the simulations, there is not much gain in the proposed kernel regression over diffusion smoothing. Although both techniques share the identical FEM discretization, the kernel regression is a parametric model while diffusion smoothing is not. The flexibility of the parametric model made us to establish the mathematical equivalence of kernel regression, diffusion smoothing and diffusion wavelets. This is the main contribution of the paper.

The method is subsequently applied in mandible growth characterization. Based on the significant directions of the growth identified in Figure 12 and 13, we observed the elongation and narrowing of the mandible, which is consistent with previous literature. For more complex growth modeling, we are currently securing additional samples.

Acknowledgment

This work was supported by NIH Research Grant R01 DC6282 (MRI and CT Studies of the Developing Vocal Tract), from the National Institute of Deafness and other Communicative Disorders (NIDCD); core grant P-30 HD03352 to the Waisman Center from the National Institute of Child Health and Human Development (NICHHD) and Vilas Associate Award from University of Wisconsin-Madison. We thank Lindell R. Gentry, Mike S. Schimek, Brian J. Whyms, Reid B. Durtschi and Dongjun Chung of University of Wisconsin-Madison for assistance with image acquisition and image segmentation.

References

Adler, R., 1981. The Geometry of Random Fields. John Wiley & Sons.

Andrade, A., Kherif, F., Mangin, J., Worsley, K., Paradis, A., Simon, O., Dehaene, S., Le Bihan, D., Poline, J.B., 2001. Detection of fMRI activation using cortical surface mapping. Human Brain Mapping 12, 79–93.

Andresen, P.R., Bookstein, F.L., Conradsen, K., Ersbøll, B.K., Marsh, J.L., Kreiborg, S., 2000. Surface-bounded growth modeling applied to human mandibles. IEEE Transactions on Medical Imaging 19, 1053–1063.

Antoine, J.P., Roşca, D., Vandergheynst, P., 2010. Wavelet transform on manifolds: old and new approaches. Applied and Computational Harmonic Analysis 28, 189–202.

Belkin, M., Niyogi, P., Sindhwani, V., 2006. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. The Journal of Machine Learning Research 7, 2399–2434.

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

Bronstein, M.M., Kokkinos, I., 2010. Scale-invariant heat kernel signatures for non-rigid shape recognition, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1704–1711.

Cachia, A., Mangin, J.F., Riviére, D., Kherif, F., Boddaert, N., Andrade, A., Papadopoulos-Orfanos, D., Poline, J.B., Bloch, I., Zilbovicius, M., Sonigo, P., Brunelle, F., Régis, J., 2003a. 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.

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

Cao, J., Worsley, K., 2001. Applications of random fields in human brain mapping. Spatial Statistics: Methodological Aspects and Applications 159, 170–182.

Chung, M., 2001. Statistical Morphometry in Neuroanatomy. Ph.D. Thesis, McGill University.

Chung, M., Hartley, R., Dalton, K., Davidson, R., 2008. Encoding cortical surface by spherical harmonics. Statistica Sinica 18, 1269–1291.

Chung, M., Robbins, S., Dalton, K., Davidson, R., Alexander, A., Evans, A., 2005. Cortical thickness analysis in autism with heat kernel smoothing. NeuroImage 25, 1256–1265.

Chung, M., Taylor, J., 2004. Diffusion smoothing on brain surface via finite element method, in: Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI), pp. 432–435.

Coifman, R., Maggioni, M., 2006. Diffusion wavelets. Applied and Computational Harmonic Analysis 21, 53–94.

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

Enlow, D.H., Hans, M.G., 1996. Essentials of Facial Growth. W.B. Saunders Company, Philadelphia.

Evans, L., 1998. Partial differential equations. American Mathematical Society.

Gaser, C., Luders, E., Thompson, P., Lee, A., Dutton, R., Geaga, J., Hayashi, K., Bellugi, U., Galaburda, A., Korenberg, J., Mills, D., Toga, A., Reiss, A., 2006. Increased local gyrification mapped in williams syndrome. NeuroImage 33, 46–54.

de Goes, F., Goldenstein, S., Velho, L., 2008. A hierarchical segmentation of articulated bodies. Computer Graphics Forum 27, 1349–1356.

Guskov, I., Wood, Z., 2001. Topological noise removal, in: Graphics Interface, pp. 19–26.

Hagler Jr., D., Saygin, A., Sereno, M., 2006. Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. NeuroImage 33, 1093–1103.

Hammond, D., Vandergheynst, P., Gribonval, R., 2011. Wavelets on graphs via spectral graph theory. Applied and Computational Harmonic Analysis 30, 129–150.

Han, J., Kim, J., Chung, C., Park, K., 2007. 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.

Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., Busa, E., Pacheco, J., Albert, M., Killiany, R., et al., 2006. Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. NeuroImage 32, 180–194.

Hernandez, V., Roman, J.E., Tomas, A., Vidal, V., 2006. A Survey of Software for Sparse Eigenvalue Problems. Technical Report STR-6. Universidad Politécnica de Valencia. Http://www.grycap.upv.es/slepc.

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

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

Joshi, S.C., Davis, B., Jomier, M., Gerig, G., 2004. Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage 23, 151–160.

Kim, S.G., Chung, M., Seo, S., Schaefer, S., van Reekum, C., Davidson, R., 2011. Heat kernel smoothing via Laplace-Beltrami eigenfunctions and its application to subcortical structure modeling, in: Pacific-Rim Symposium on Image and Video Technology (PSIVT). Lecture Notes in Computer Science (LNCS), pp. 36–47.

Kim, W., Pachauri, D., Hatt, C., Chung, M., Johnson, S., Singh, V., 2012. Wavelet based multi-scale shape features on arbitrary surfaces for cortical thickness discrimination, in: Advances in Neural Information Processing Systems, pp. 1250–1258.

Lai, Z., Hu, J., Liu, C., Taimouri, V., Pai, D., Zhu, J., Xu, J., Hua, J., 2010. Intra-patient supine-prone colon registration in CT colonography using shape spectrum, in: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010, pp. 332–339.

Lehoucq, R., Sorensen, D., Yang, C., 1998. ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods. SIAM Publications, Philadelphia.

Lerch, J.P., Evans, A., 2005. Cortical thickness analysis examined through power analysis and a population simulation. NeuroImage 24, 163–173.

Lévy, B., 2006. Laplace-Beltrami eigenfunctions towards an algorithm that "understands" geometry, in: IEEE International Conference on Shape Modeling and Applications, p. 13.

Lewis, A., Roche, A., Wagner, B., 1982. Growth of the mandible during pubescence. The Angle Orthodontist 52, 325–342.

Loubele, M., Maes, F., Schutyser, F., Marchal, G., Jacobs, R., Suetens, P., 2006. Assessment of bone segmentation quality of cone-beam CT versus multislice spiral CT: a pilot study. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology 102, 225–234.

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

Luders, E., Thompson, P., Narr, K., Toga, A., Jancke, L., Gaser, C., 2006b. A curvature-based approach to estimate local gyrification on the cortical surface. NeuroImage 29, 1224–1230.

Mahadevan, S., Maggioni, M., 2006. Value function approximation with diffusion wavelets and laplacian eigenfunctions. Advances in neural information processing systems 18, 843.

Malladi, R., Ravve, I., 2002. Fast difference schemes for edge enhancing Beltrami flow, in: Proceedings of Computer Vision-ECCV, Lecture Notes in Computer Science (LNCS), pp. 343–357.

Miller, M., Qiu, A., 2009. The emerging discipline of computational functional anatomy. Neuroimage 45, S16–S39.

Nain, D., Styner, M., Niethammer, M., Levitt, J., Shenton, M., Gerig, G., Bobick, A., Tannenbaum, A., 2007. Statistical shape analysis of brain structures using spherical wavelets, in: IEEE Symposium on Biomedical Imaging ISBI.

Niethammer, M., Reuter, M., Wolter, F., Bouix, S., Peinecke, N., Koo, M., Shenton, M., 2007. Global Medical Shape Analysis Using the Laplace-Beltrami Spectrum. Lecture Notes in Computer Science 4791, 850.

Nilsson, J., Sha, F., Jordan, M., 2007. Regression on manifolds using kernel dimension reduction, in: Proceedings of the 24th international conference on Machine learning, ACM. pp. 697–704.

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

Qiu, A., Bitouk, D., Miller, M., 2006. Smooth functional and structural maps on the neocortex via orthonormal bases of the laplace-beltrami operator. IEEE Transactions on Medical Imaging 25, 1296–1396.

Qiu, A., Miller, M., 2008. Multi-structure network shape analysis via normal surface momentum maps. NeuroImage 42, 1430–1438.

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

Reuter, M., Wolter, F.E., Peinecke, N., 2006. Laplace-Beltrami spectra as shape-dna of surfaces and solids. Computer-Aided Design 38, 342–366.

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

Rosenberg, S., 1997. The Laplacian on a Riemannian Manifold. Cambridge University Press.

Rustamov, R., 2007. Laplace-Beltrami eigenfunctions for deformation invariant shape representation, in: Proceedings of the fifth Eurographics symposium on Geometry processing, p. 233.

Schölkopf, B., Smola, A.J., 2002. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press.

Scott, J.H., 1976. Dento-facial Development and Growth. Pergamon Press, Oxford.

Seo, S., Chung, M., 2011. Laplace-Beltrami eigenfunction expansion of cortical manifolds, in: Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, IEEE. pp. 372–375.

Seo, S., Chung, M., Voperian, H., 2011. Mandible shape modeling using the second eigenfunction of the laplace-beltrami operator, in: Proceedings of SPIE Medical Imaging, p. 79620Z.

Seo, S., Chung, M., Vorperian, H., 2010. Heat kernel smothing using laplace-beltrami eigenfunctions, in: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010, pp. 505–512.

Shawe-Taylor, J., Cristianini, N., 2004. Kernel methods for pattern analysis. Cambridge University Press.

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

Shen, L., Saykin, A., Chung, M., Huang, H., Ford, J., Makedon, F., McHugh, T., Rhodes, C., 2006. 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.

Shi, Y., Dinov, I., Toga, A.W., 2009. Cortical shape analysis in the Laplace-Beltrami feature space, in: 12th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2009), pp. 208–215.

Shi, Y., Lai, R., Kern, K., Sicotte, N., Dinov, I., Toga, A.W., 2008. Harmonic surface mapping with Laplace-Beltrami eigenmaps, in: 11th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2008), pp. 147–154.

Smartt Jr., J.M., Low, D.W., Bartlett, S.P., 2005. The pediatric mandible: I. A primer on growth and development. Plast. Reconstr. Surg. 116, 14e–23e.

Sochen, N., Kimmel, R., Malladi, R., 1998. A general framework for low level vision. IEEE Transactions on Image Processing 7, 310–318.

Steinke, F., Hein, M., 2008. Non-parametric regression between manifolds. Advances in Neural Information Processing Systems .

Stratemann, S.A., Huang, J.C., Maki, K., Hatcher, D.C., Miller, A.J., 2010. Evaluating the mandible with cone-beam computed tomography. American Journal of Orthodontics and Dentofacial Orthopedics 137, S58–S70.

Styner, M., Oguz, I., Xu, S., Brechbuhler, C., Pantazis, D., Levitt, J., Shenton, M., Gerig, G., 2006. Framework for the statistical shape analysis of brain structures using spharm-pdm, in: Insight Journal, Special Edition on the Open Science Workshop at MICCAI.

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

Tang, B., Sapiro, G., Caselles, V., 1999. Direction diffusion, in: The Proceedings of the Seventh IEEE International Conference on Computer Vision, pp. 2:1245–1252.

Tasdizen, T., Whitaker, R., Burchard, P., Osher, S., 2006. Geometric surface smoothing via anisotropic diffusion of normals, in: Geometric Modeling and Processing, pp. 687–693.

Taubin, G., 2000. Geometric Signal Processing on Polygonal Meshes, in: EUROGRAPHICS.

Taylor, J., Worsley, K., 2007. Detecting sparse signals in random fields, with an application to brain mapping. Journal of the American Statistical Association 102, 913–928.

Vaillant, M., Qiu, A., Glaunès, J., Miller, M., 2007. Diffeomorphic metric surface mapping in subregion of the superior temporal gyrus. NeuroImage 34, 1149–1159.

Vallet, B., Lévy, B., 2008. Spectral geometry processing with manifold harmonics. Computer Graphics Forum 27, 251–260.

Van Den Boomgaard, R., Van Balen, R., 1992. Methods for fast morphological image transforms using bitmapped binary images. CVGIP: Graphical Models and Image Processing 54, 252–258.

Walker, G.F., Kowalski, C.J., 1972. On the growth of the mandible. American Journal of Physical Anthropology 36, 111–118.

Wang, F.Y., 1997. Sharp explict lower bounds of heat kernels. Annals of Probability 24, 1995–2006.

Wang, Y., Zhang, J., Gutman, B., Chan, T., Becker, J., Aizenstein, H., Lopez, O., Tamburo, R., Toga, A., Thompson, P., 2010. Multivariate tensor-based morphometry on surfaces: Application to mapping ventricular abnormalities in HIV/AIDS. NeuroImage 49, 2141–2157.

Wardetzky, M., 2008. Convergence of the cotangent formula: An overview, in: Bobenko, A., Sullivan, J., Schröder, P., Ziegler, G. (Eds.), Discrete Differential Geometry. Birkhäuser Basel. volume 38 of *Oberwolfach Seminars*, pp. 275–286.

Wood, Z., Hoppe, H., Desbrun, M., Schröder, P., 2004. Removing excess topology from isosurfaces. ACM Transactions on Graphics (TOG) 23, 190–208.

Worlsey, K., Poline, J.B., Vandal, A., Friston, K., 1995. Test for distributed, non-focal brain activations. NeuroImage 2, 173–181.

Worsley, K., 2003. Detecting activation in fMRI data. Statistical Methods in Medical Research. 12, 401–418.

Worsley, K., Taylor, J., Tomaiuolo, F., Lerch, J., 2004. Unified univariate and multivariate random field theory. NeuroImage 23, S189–195.

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

Yger, F., Rakotomamonjy, A., 2011. Wavelet kernel learning. Pattern Recognition 44, 2614–2629.

Yotter, R.A., Dahnke, R., Gaser, C., 2009. Topological correction of brain surface meshes using spherical harmonics, in: Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer. pp. 125–132.

Yushkevich, P., Zhang, H., Simon, T., Gee, J., 2008. Structure-specific statistical mapping of white matter tracts. NeuroImage 41, 448–461.

Zhang, H., van Kaick, O., Dyer, R., 2007. Spectral methods for mesh processing and analysis, in: EUROGRAPHICS, pp. 1–22.

Zhang, H., van Kaick, O., Dyer, R., 2010. Spectral mesh processing. Computer Graphics Forum 29, 1865–1894.

Zhong, J., Qiu, A., 2010. Multi-manifold diffeomorphic metric mapping for aligning cortical hemispheric surfaces. Neuroimage 49, 355–365.

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