

Heat Kernel Smoothing on Manifolds and Its Application to Hyoid Bone Growth Modeling



Moo K. Chung, Nagesh Adluru, and Hourì K. Vorperian

Abstract We present a unified heat kernel smoothing framework for modeling 3D anatomical surface data extracted from medical images. Due to image acquisition and preprocessing noises, it is expected the medical imaging data is noisy. The surface data of the anatomical structures is regressed using the weighted linear combination of Laplace-Beltrami (LB) eigenfunctions to smooth out noisy data and perform statistical analysis. The method is applied in characterizing the 3D growth pattern of human hyoid bone between ages 0 and 20 obtained from CT images. We detected a significant age effect on localized parts of the hyoid bone.

Keywords Heat kernel smoothing · Hyoid bone growth · Random field theory · Laplace Beltrami eigenfunctions · Diffusion on manifolds

1 Introduction

For normally developing children, age and sex could be major factors that affect the structure and function of growing hyoid bone. As in other developmental studies [22, 55, 56], we expect highly localized complex growth pattern to emerge between

M. K. Chung (✉)

Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison, WI, USA

e-mail: mkchung@wisc.edu

N. Adluru

Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin, Madison, WI, USA

e-mail: adluru@wisc.edu

H. K. Vorperian

Vocal Tract Development Laboratory, Waisman Center, University of Wisconsin, Madison, WI, USA

e-mail: vorperian@waisman.wisc.edu

© Springer Nature Switzerland AG 2020

Y. Zhao, D.-G. (Din) Chen (eds.), *Statistical Modeling in Biomedical Research*, Emerging Topics in Statistics and Biostatistics, https://doi.org/10.1007/978-3-030-33416-1_12

235

ages 0 and 20 in the hyoid bone. Growth is expected to extend laterally with respect to the surface of the bone. However, it is unclear what specific parts of the hyoid bone are growing. This provides a biological motivation for a need to develop a local surface-based morphometric technique beyond simple volumetric techniques that cannot detect localized subtle anatomical changes along the surface of the hyoid bone composting of three segments—a central hyoid body with two greater cornua (horns) [11, 21].

The end results of existing surface-based morphometric studies in medical imaging are statistical parametric maps (SPM) that show the statistical significance of growth at each surface mesh vertex [22, 48, 64]. In order to obtain stable and robust SPM, various signal smoothing and filtering methods have been proposed. Among them, diffusion equations, kernel smoothing, and wavelet-based approaches are probably the most popular. Diffusion equations have been widely used in image processing as a form of noise reduction starting with Perona and Malik in 1990s [46]. Although numerous techniques have been developed for performing diffusion along surfaces [2, 21, 22, 43, 51–53], many approaches are nonparametric and requires the finite element or finite difference schemes which are known to suffer various numerical instabilities [16, 18].

Recently, few regression models are proposed on manifolds. In [40], Laplace-Beltrami operator based functional principal component analysis was proposed. In [25], Fréchet mean based regression model was proposed on manifolds. Kernel smoothing based models have been also proposed for surface and manifolds data [6, 17, 18]. The kernel methods basically smooth data as the weighted average of neighboring mesh vertices using mostly a Gaussian kernel and its iterative application is supposed to approximate the diffusion process. Recently, wavelets have been popularized for surface and graph data [33, 36, 38]. Spherical wavelets have been used on brain surface data that has been mapped onto a sphere [8, 44]. Since wavelet basis functions have local support in both space and scale, the wavelet coefficients from the scale-space decomposition using the spherical wavelets provide shape features that describe local shape variation at a variety of scales and spatial locations. However, spherical wavelets have an intrinsic problem that they require to establish a smooth mapping from the surface to a unit sphere, which introduces a serious metric distortion. The spherical mapping such as conformal mapping introduces serious metric distortion which usually compounds SPM [28, 34]. Furthermore, such basis functions defined on a sphere seem to be suboptimal than those directly defined on anatomical surfaces in detecting locations or scales of shape variations. To remedy the limitation of the spherical wavelets, the spectral graph wavelet transform defined on a graph has been applied to arbitrary surface meshes by treating surface meshes as graphs [3, 29, 38]. The wavelet transform is a powerful tool decomposing a signal or function into a collection of components localized at both location and scale. Although all three methods (diffusion-, kernel- and wavelet-based) look different from each other, it is possible to develop a unified framework that relates all of them in a coherent mathematical framework [16].

Starting with a symmetric positive definite kernel, we propose a unified kernel smoothing framework within the Hilbert space theory [23]. The proposed kernel

smoothing works for any symmetric positive definite kernel, which behaves like weights between two functional data. We show how this facilitates a coherent statistical inference for functional signals defined on an arbitrary manifold. The focus of this paper is on the development of the proposed kernel smoothing on manifolds.

The structure of this paper is as follows. First, we present a unified bivariate kernel smoothing that is related to diffusion-like equations on manifolds. The proposed kernel regression inherits various mathematical and statistical properties of diffusion-like equations. Then, we show the relationship between the kernel smoothing and recently popular spectral graph wavelets for manifolds. The proposed kernel smoothing is shown to be equivalent to the wavelet transform. This mathematical equivalence levitates a need for constructing wavelets using a complicated computational machinery as often done in previous diffusion wavelet constructions [3, 29, 36, 38]. A unified statistical inference framework is then developed for the kernel method via Worsley's random field theory [54, 63]. This levitates the need for using time consuming nonparametric procedures such as false discovery rates (FDR) [7, 27] or permutation tests [9, 15, 20, 31] that do not have explicate control over the scale and smoothness of models. Finally, we illustrate how the kernel smoothing procedure can be used to localize the disconnected hyoid bone growth pattern in human.

2 Preliminary

Let us illustrate two statistical problems in the Euclidean space that motivate the development of the proposed kernel smoothing on manifolds. Consider measurements f_i sampled at point $p_i \in \mathbb{R}^d$. The measurements are usually modeled as

$$f_i = h(p_i) + \epsilon_i$$

with mean zero noise ϵ_i and unknown mean function h that has to be estimated. In the traditional kernel regression framework [6, 24, 45], the mean function h is estimated in the weighted least squares fashion:

$$\hat{h}(p) = \sum_{j=1}^k G(p - p_i) f_i,$$

where G is given by Nadaraya-Watson type of normalized kernels. In the local polynomial regression framework [24], h is estimated as

$$\hat{h}(p) = \arg \min_{\beta_0, \dots, \beta_k} \sum_{i=1}^n G(p - p_i) \left| f_i - \sum_{j=0}^k \beta_j (p - p_i)^j \right|^2. \quad (1)$$

In many related local polynomial or kernel regression frameworks, kernel G and polynomial basis $\{p^j\}$ are translated by the amount of p_i in fitting the data locally. In this fashion, at each data point p_i , exactly the same shape of kernel and distance can be used. However, one immediately encounters a difficulty of directly generalizing the Euclidean formulation (1) to an arbitrary surface since it is unclear how to translate the kernel and basis in a coherent fashion. To remedy this problem, many recent kernel regression frameworks on manifolds use bivariate kernel $G(p, q)$ and bypass the problem of translating a univariate kernel [6]. By simply changing the second argument, it has the effect of translating the kernel.

A similar problem is also encountered in wavelets in the Euclidean space. Consider wavelet basis $W_{t,q}(p)$ obtained from a mother wavelet W with scale parameter t and translation parameter q :

$$W_{t,q}(p) = \frac{1}{t} W\left(\frac{p-q}{t}\right). \quad (2)$$

Scaling a function on a surface is trivial. But the difficulty arises when one tries to define a mother wavelet and translate it on a surface. It is not straightforward to generalize the Euclidean formulation (2) to an arbitrary manifold. If one tries to modify the existing spherical wavelets to an arbitrary surface [8, 44], one also encounters the lack of regular grids on the surface. The recent work based on the spectral graph wavelet transform bypass this problem by also taking a bivariate kernel as a mother wavelet [3, 29, 38, 42]. To remedy these two different but related problems, we propose to use a bivariate kernel and bypass the problem of translating a univariate kernel. By simply changing the second argument, it has the effect of translating the kernel.

3 Methods

In many anatomical surface studies in medical imaging, measurements are sampled densely at each mesh vertex so it is more practical to model the measurements as smooth functions. Consider a functional measurement f defined on a manifold $\mathcal{M} \subset \mathbb{R}^d$. We assume the following additive model:

$$f(p) = h(p) + \epsilon(p), \quad (3)$$

where h is the unknown signal to be estimated and ϵ is a zero-mean random field, possibly Gaussian. The manifold \mathcal{M} can be a single connected component or multiple disjoint components as our hyoid bone application (Fig. 1). We further assume $f \in L^2(\mathcal{M})$, the space of square integrable functions on \mathcal{M} with the inner product

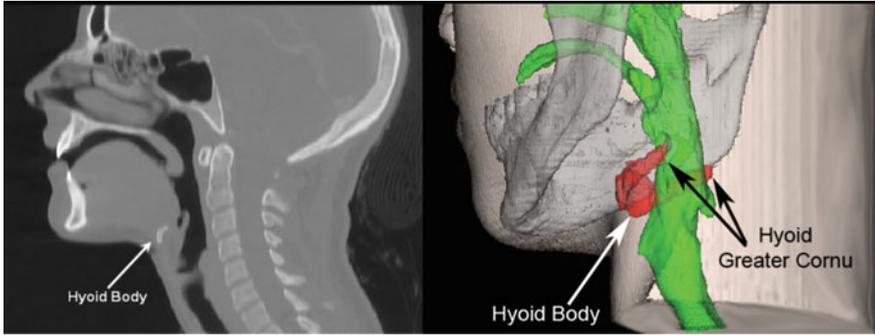


Fig. 1 CT image showing the location of the hyoid bone (left), and 3D model (right) showing the relative location of the hyoid bone (red) with respect to the mandible (gray) and vocal tract (green)

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

where μ is the Lebesgue measure. $\mu(\mathcal{M})$ will measure the volume of \mathcal{M} in d -dimension [11, 16]. Define a self-adjoint operator \mathcal{L} satisfying

$$\langle g_1, \mathcal{L}g_2 \rangle = \langle \mathcal{L}g_1, g_2 \rangle$$

for all $g_1, g_2 \in L^2(\mathcal{M})$. Then \mathcal{L} induces the eigenvalues λ_j and eigenfunctions ψ_j on \mathcal{M} (Fig. 2):

$$\mathcal{L}\psi_j = \lambda_j\psi_j. \tag{4}$$

Without loss of generality, we can order the eigenvalues

$$0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \dots$$

We can show that the eigenfunctions ψ_j form an orthonormal basis in $L^2(\mathcal{M})$. We will consider a smooth symmetric positive definite kernel of the form

$$K(p, q) = \sum_{j=0}^{\infty} \tau_j \psi_j(p)\psi_j(q) \tag{5}$$

for some τ_j in this paper. The constants τ_j are identified as follows. Apply the kernel convolution on the eigenfunction ψ_j :

$$K * \psi_j(p) = \int_{\mathcal{M}} K(p, q)\psi_j(q) d\mu(q). \tag{6}$$

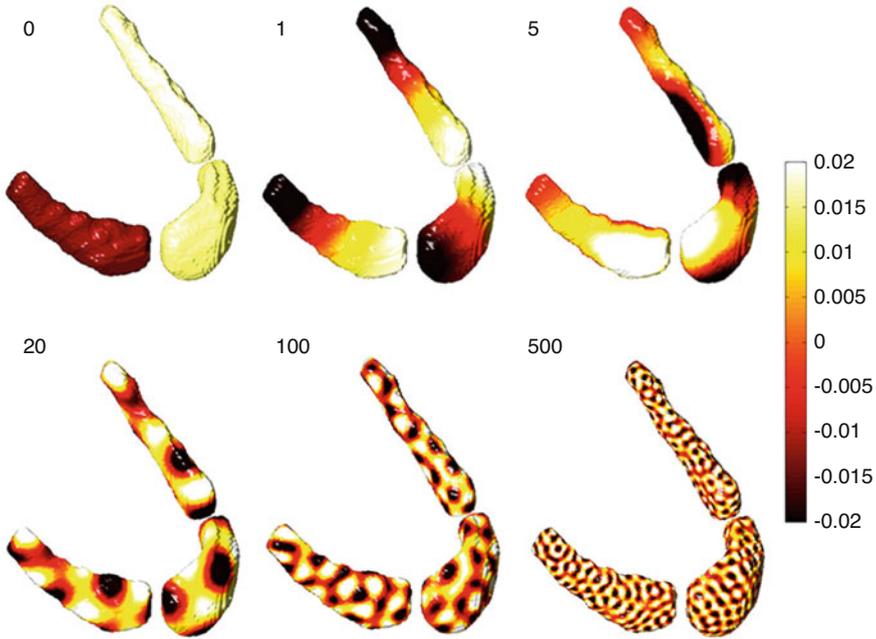


Fig. 2 Laplace-Beltrami eigenfunctions ψ_j of various degrees ($j = 0, 1, 5, 20, 100, 500$) on the template. The first eigenfunction is constant in each component. As the degree increases, the spatial frequency increases

Substituting (5) into (6), we have

$$K * \psi_j(p) = \tau_j \psi_j(p)$$

indicating τ_j and ψ_j must be the eigenvalues and eigenfunctions of the convolution (6). Note ψ_j are eigenfunctions of self-adjoint operator \mathcal{L} and kernel convolution at the same time.

Example 1 For $\tau_j = e^{-\lambda t}$, we have heat kernel

$$K(p, q) = \sum_{j=0}^{\infty} e^{-\lambda t} \psi_j(p) \psi_j(q), \tag{7}$$

where t is the bandwidth of kernel. The heat kernel has been often used in numerous studies but without much theoretical justification [16, 32, 37, 49]. For this study, we will denote the heat kernel as $H_t(p, q)$ to explicitly show that the spread of the kernel is determined by diffusion time t [17, 18].

3.1 Kernel Smoothing on Manifolds

Consider subspace $\mathcal{H}_k \subset L^2(\mathcal{M})$ spanned by the orthonormal basis $\{\psi_j\}$, i.e.,

$$\mathcal{H}_k = \left\{ \sum_{j=0}^k \beta_j \psi_j(p) : \beta_j \in \mathbb{R} \right\}.$$

Then the least squares estimation (LSE) of h in \mathcal{H}_k is given by the shortest distance from f to \mathcal{H}_k [14, 16]:

$$\widehat{h}(p) = \arg \min_{h \in \mathcal{H}_k} \int_{\mathcal{M}} |f(p) - h(p)|^2 d\mu(p) = \sum_{j=0}^k f_j \psi_j(p), \tag{8}$$

where $f_j = \langle f, \psi_j \rangle$ are the Fourier coefficients. Figure 3 shows an example of LSE with \mathcal{L} as the Laplace-Beltrami operator and $k = 1000$. This is a special case of Fourier series expansion that tends to suffer the Gibbs phenomenon, i.e., ringing artifact [13, 26]. The Gibbs phenomenon can be effectively removed if the Fourier series expansion converges fast enough as the number of basis functions goes to infinity. By weighting the Fourier coefficients exponentially smaller, we can make

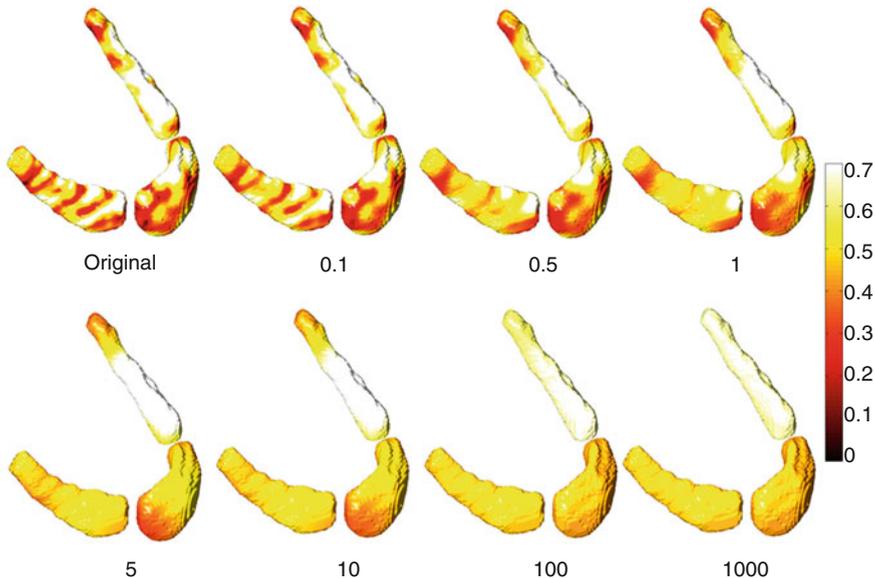


Fig. 3 Heat kernel smoothing using different bandwidth between 0.1 and 1000. As the bandwidth increases, the kernel regression becomes inversely proportional to the square root of the surface area

the representation converges faster; this can be achieved by additionally weighting the squared residuals in Eq. (8) with some weights. Thus, we propose to estimate h by minimizing the weighted distance to the space \mathcal{H}_k :

$$\widehat{h}(p) = \arg \min_{h \in \mathcal{H}_k} \int_{\mathcal{M}} \int_{\mathcal{M}} K(p, q) |f(q) - h(p)|^2 d\mu(q) d\mu(p). \quad (9)$$

Without loss of generality, we will assume the kernel to be a probability distribution

$$\int_{\mathcal{M}} K(p, q) d\mu(q) = 1$$

for all $p \in \mathcal{M}$. The solution of (9) has the following analytic expression.

Theorem 1

$$\widehat{h}(p) = \arg \min_{h \in \mathcal{H}_k} \int_{\mathcal{M}} \int_{\mathcal{M}} K(p, q) |f(q) - h(p)|^2 d\mu(q) d\mu(p) = \sum_{j=0}^k \tau_j f_j \psi_j,$$

where $f_j = \langle f, \psi_j \rangle$ are Fourier coefficients.

Proof Any function $h \in \mathcal{H}_k$ can be expressed as

$$h(p) = \sum_{j=0}^k \beta_j \psi_j(p). \quad (10)$$

Then by plugging (10) into the inner integral $I(p)$, it becomes

$$I(p) = \int_{\mathcal{M}} K(p, q) \left| f(q) - \sum_{j=0}^k \beta_j \psi_j(p) \right|^2 d\mu(q).$$

Simplifying the expression, we obtain

$$I(p) = \sum_{j=0}^k \sum_{j'=0}^k \psi_j(p) \psi_{j'}(p) \beta_j \beta_{j'} - 2K * f(p) \sum_{j=0}^k \psi_j(p) \beta_j + K * f^2(p). \quad (11)$$

The kernel can be written as

$$K(p, q) = \sum_{j'=0}^{\infty} \tau_{j'} \psi_{j'}(p) \psi_{j'}(q). \quad (12)$$

The convolution is then written as

$$K * f(p) = \sum_{j'=0}^{\infty} \tau_{j'} f_{j'} \psi_{j'}(p).$$

Since I is an unconstrained positive semidefinite quadratic program (QP) in β_j , there is no unique global minimizer of I without additional linear constraints. Integrating I further with respect to $d\mu(p)$, we collapse (11) to a positive definite QP, which yields a unique global minimizer:

$$\int_{\mathcal{M}} I(p) d\mu(p) = \sum_{j=0}^k \beta_j^2 - 2 \sum_{j=0}^k \tau_j f_j \beta_j + \text{const.}$$

The minimum of the above integral is obtained when all the partial derivatives with respect to β_j vanish, i.e.

$$\int_{\mathcal{M}} \frac{\partial I}{\partial \beta_j} d\mu(p) = 2\beta_j - 2\tau_j f_j = 0$$

for all j . Hence $\sum_{j=0}^k \tau_j f_j \psi_j$ must be the unique minimizer. \square

Theorem 1 generalizes the weighted spherical harmonic (SPHARM) representation on a unit sphere to an arbitrary manifold [14]. Theorem 1 implies that the kernel regression can be performed by simply computing the Fourier coefficients $f_j = \langle f, \psi_j \rangle$ without doing any numerical optimization. The numerically difficult optimization problem is then reduced to the problem of computing Fourier coefficients. If the kernel K is the Dirac-delta function, the kernel regression simply collapses to the least squares estimation (LSE) which results in the standard Fourier series, i.e.

$$\widehat{h}(p) = \arg \min_{h \in \mathcal{H}_k} \int_{\mathcal{M}} |f(q) - h(q)|^2 d\mu(q) = \sum_{j=0}^k f_j \psi_j.$$

It can be also shown that as $k \rightarrow \infty$, the kernel regression

$$\widehat{h} = \sum_{j=0}^k \tau_j f_j \psi_j$$

converges to convolution $K * f$ establishing the connection to the manifold-based kernel smoothing framework [5, 18]. Hence, asymptotically the proposed kernel regression should inherit many statistical properties of the usual kernel smoothing.

3.2 Properties of Kernel Smoothing

Kernel smoothing can be shown to be related to the following diffusion-like Cauchy problem [13, 14].

Theorem 2 For an arbitrary self-adjoint differential operator \mathcal{L} , the unique solution of the following initial value problem

$$\frac{\partial g(p, t)}{\partial t} + \mathcal{L}g(p, t) = 0, g(p, t = 0) = f(p) \quad (13)$$

is given by

$$g(p, t) = \sum_{j=0}^{\infty} e^{-\lambda_j t} f_j \psi_j(p). \quad (14)$$

Proof For each fixed t , $g(p, t)$ can be written as

$$g(p, t) = \sum_{j=0}^{\infty} c_j(t) \psi_j(p). \quad (15)$$

Then

$$\mathcal{L}g(p, t) = \sum_{j=0}^{\infty} c_j(t) \lambda_j \psi_j(p). \quad (16)$$

Substituting (15) and (16) into (13), we obtain

$$\frac{\partial c_j(t)}{\partial t} + \lambda_j c_j(t) = 0 \quad (17)$$

for all j . The solution of equation (17) is given by $c_j(t) = b_j e^{-\lambda_j t}$. So we have a solution

$$g(p, t) = \sum_{j=0}^{\infty} b_j e^{-\lambda_j t} \psi_j(p).$$

At $t = 0$, we have

$$g(p, 0) = \sum_{j=0}^{\infty} b_j \psi_j(p) = f(p).$$

The coefficients b_j must be the Fourier coefficients, i.e.,

$$b_j = \langle f, \psi_j \rangle = f_j.$$

□

For a particular choice of kernel K with $\tau_j = e^{-\lambda_j t}$, the proposed kernel regression $\hat{h} = \sum_{j=0}^k \tau_j f_j \psi_j$ should converge to the solution of the diffusion-like equation.

Example 2 If \mathcal{L} is the Laplace-Beltrami operator, (13) becomes an isotropic diffusion equation as a special case and we are then dealing with heat kernel

$$H_t(p, q) = \sum_{j=0}^{\infty} e^{-\lambda_j t} \psi_j(p) \psi_j(q),$$

which is often explored mathematical objects in various areas [5, 18].

In order to construct wavelets on an arbitrary graph and mesh, diffusion wavelet transform has been proposed recently [3, 29, 38]. The diffusion wavelet construction has been fairly involving so far. However, its mathematical structure is related to the proposed kernel smoothing. For scale function g that satisfies the admissibility conditions [3, 29, 36, 38], diffusion wavelet $W_{t,p}(p)$ at position p and scale t is given by

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

If we let $\tau_j = g(\lambda_j t)$, the diffusion wavelet transform is given by

$$\langle W_{t,p}, f \rangle = \int_{\mathcal{M}} W_{t,q}(p) f(p) d\mu(p) = \sum_{j=0}^k \tau_j f_j \psi_j(q),$$

which is the exactly kernel smoothing we introduced. Hence, the diffusion wavelet transform can be simply obtained by doing the kernel smoothing with specific scale function g [38]. If we let $g(\lambda_j t) = e^{-\lambda_j t}$, we have

$$W_{t,p}(q) = H_t(p, q),$$

which is a heat kernel. The bandwidth t of heat kernel controls resolution while the translation is done by shifting one argument in the kernel. Thus, although heat kernel smoothing is not exactly diffusion wavelet, it shares the same algebraic formalism and behaves similarly. Although the kernel smoothing is constructed using global basis functions ψ_j , the kernel regression at each point p coincides with the diffusion wavelet transform at that point. Hence, just like wavelets, the kernel smoothing will have the localization property of wavelets.

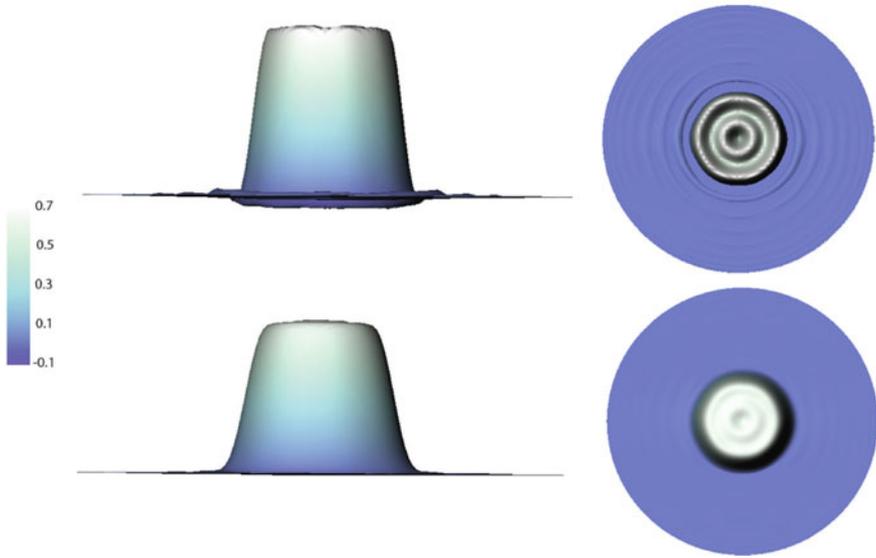


Fig. 4 The Gibbs phenomenon on a hat shaped simulated surface showing the ringing effect on the SPHARM expansion (top) and the reduced effect on the heat kernel smoothing (bottom) [12, 13]. 7225 basis functions were used for the both cases and the bandwidth $t = 0.001$ is used for heat kernel smoothing

Another important property of heat kernel smoothing is the ability to reduce the Gibbs phenomenon, which often occurs when we tried to represent signals with rapid changes [12–14]. Example 3 illustrates how heat kernel smoothing can be use in reducing ringing artifacts in a 3D step function.

Example 3 A hat-shaped step function is simulated in 3D as $z = 1$ for $x^2 + y^2 < 1$ and $z = 0$ for $1 \leq x^2 + y^2 \leq 2$ (Fig. 4). Then the step function is reconstructed using the SPHARM expansion via LSE (top) and kernel regression (bottom). In the both cases, up to 7225 basis functions were used. For the kernel regression, the heat kernel with bandwidth $t = 0.0001$ is used. LSE clearly shows the visible Gibbs phenomenon, i.e., ringing artifact [13, 26] compared to the kernel regression.

3.3 Numerical Implementation

In this study, the Laplace-Beltrami operator is chosen as the self-adjoint operators \mathcal{L} of choice. The eigenfunctions of the Laplace-Beltrami operator on an arbitrary curved surface is analytically unknown. So it is necessary to discretize (4) using the Cotan formulation as a generalized eigenvalue problem [19, 47, 66]:

$$\mathbf{C}\psi = \lambda\mathbf{A}\psi, \quad (18)$$

where \mathbf{C} is the stiffness matrix, \mathbf{A} is the mass matrix and $\boldsymbol{\psi} = (\psi(p_1), \dots, \psi(p_n))'$ is the eigenfunction evaluated at n mesh vertices. Once we obtained the basis functions ψ_j , the corresponding LB-eigenfunction expansion coefficients β_j are estimated as

$$\beta_j = \mathbf{f}' \mathbf{A} \boldsymbol{\psi}_j,$$

where $\mathbf{f} = (f(p_1), \dots, f(p_n))'$ and $\boldsymbol{\psi}_j = (\psi_j(p_1), \dots, \psi_j(p_n))'$ [66]. Figure 2 shows few representative LB-eigenfunctions on the hyoid surface. For heat kernel smoothing, we used the bandwidth $t = 5$ and 500 LB-eigenfunctions on the surface of the hyoid bone. The number of eigenfunctions used is more than sufficient to guarantee relative error less than 0.3% in our data. The MATLAB code for computing the eigenfunctions and performing heat kernel smoothing is available at <http://www.stat.wisc.edu/~mchung/mandible>.

3.4 Statistical Inference

We are interested in determining the significance of functional signals on manifolds. We borrow the statistical parametric mapping (SPM) framework for analyzing and visualizing statistical tests on surfaces that is often used in brain image analysis [2, 17, 39, 57, 62, 65]. Since test statistics are constructed over all mesh vertices on the surfaces, the multiple comparisons correction is needed. For continuous functional data, the random field theory is often used [54, 62, 63]. The random field theory assumes the measurements to be a smooth Gaussian random field. Heat kernel smoothing will make the data more smooth and Gaussian and increase the signal-to-noise ratio [17].

Consider a functional measurements f_1, \dots, f_n on manifold \mathcal{M} . In the simplest statistical setting, the measurements can be modeled as

$$f_i(p) = h(p) + \epsilon_i(p),$$

where h is an unknown group level signal and ϵ_i is a zero-mean Gaussian random field [63]. At each fixed point p , we are assuming $\epsilon_i \sim N(0, \sigma^2)$.

We are interested in determining the significance of h , i.e.

$$H_0 : h(p) = 0 \text{ for all } p \in \mathcal{M} \text{ vs. } H_1 : h(p) > 0 \text{ for some } p \in \mathcal{M}. \quad (19)$$

Note that any point p_0 that gives $h(p_0) > 0$ is considered as signal. The hypothesis (19) is 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 smoothing. Subsequently, a test statistic is given by a T-field $T(p)$ or a F-field, which is simply given by the square of the T-field [62, 63].

Under H_0 , the type-I error of testing hypotheses (19) is given by

$$\begin{aligned} \alpha &= P(T(p) > z \text{ for some } p \in \mathcal{M}) \\ &= 1 - P(T(p) \leq z \text{ for all } p \in \mathcal{M}) \\ &= 1 - P\left(\sup_{p \in \mathcal{M}} T(p) \leq z\right) \\ &= P\left(\sup_{p \in \mathcal{M}} T(p) > z\right) \end{aligned}$$

for observed threshold z , which is the maximum $T(p)$ in the whole region \mathcal{M} . Note we are taking the sup operator over all \mathcal{M} . For sufficiently high threshold z , the multiple comparisons corrected type-I error of testing hypothesis (19) is given by

$$P\left(\sup_{p \in \mathcal{M}} T(p) > z\right) = \sum_{j=0}^d \mu_j(\mathcal{M}) \rho_j(z),$$

where $\mu_d(\mathcal{M})$ is the j -th Minkowski functional or intrinsic volume of \mathcal{M} and ρ_j is the j -th Euler characteristic (EC) density of T-field [1, 54, 59, 63]. Since the hyoid bone is compact with no boundary but has three disconnected components, the Minkowski functionals are simply

$$\begin{aligned} \mu_2(\mathcal{M}) &= \text{area}(\mathcal{M})/2 \\ \mu_1(\mathcal{M}) &= 0 \\ \mu_0(\mathcal{M}) &= \chi(\mathcal{M}) = 3 \times 2. \end{aligned}$$

The term μ_1 is zero since there is no boundary and μ_0 is simply the Euler characteristic of the template surface. Note that the Euler characteristic of a closed surface with no hole or handle is 2 and there are three such surfaces. The EC-densities of the T-field with ν degrees of freedom is given by

$$\begin{aligned} \rho_0(z) &= 1 - P(T_\nu \leq z), \\ \rho_1(z) &= \frac{1}{\sqrt{2t^2}} \cdot \frac{1}{2\pi} \left(1 + \frac{z^2}{\nu}\right)^{-(\nu-1)/2}, \\ \rho_2(z) &= \frac{1}{2t^2} \cdot \frac{1}{(2\pi)^{3/2}} \frac{\Gamma(\frac{\nu+1}{2})}{(\frac{\nu}{2})^{1/2} \Gamma(\frac{\nu}{2})} z \left(1 + \frac{z^2}{\nu}\right)^{-(\nu-1)/2}. \end{aligned}$$

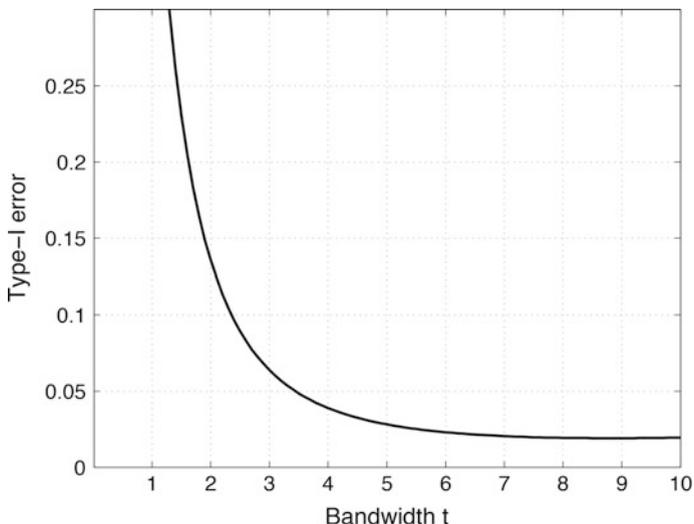


Fig. 5 The type-I error plot over bandwidth t of kernel smoothing for testing the difference between the groups I and III on the middle hyoid bone. As the bandwidth increases, the multiple comparisons corrected type-I error decreases. The bandwidth 5 is chosen for the study. The choice of the bandwidth around 5 does not change the over-all type-I error much

The EC-density of the F-field is similarly given in [54, 63]. The EC-density has the kernel bandwidth t in the formulation so the inference is done at a particular smoothing scale. Figure 5 shows the type-I error plot over different bandwidth t of the kernel regression in our application. As the bandwidth t goes to zero, the type-I error increases. When $t = 0$, the kernel regression collapse to the usual Fourier series expansion. Note that the LB-eigenfunction expansion with 500 eigenfunctions is close to the original data without any smoothing. Hence, the proposed kernel smoothing can be viewed as having substantially smaller type-I error compared to the LB-eigenfunction expansion and the original data demonstrating a better statistical performance. The type-II error and the statistical power can be similarly computed.

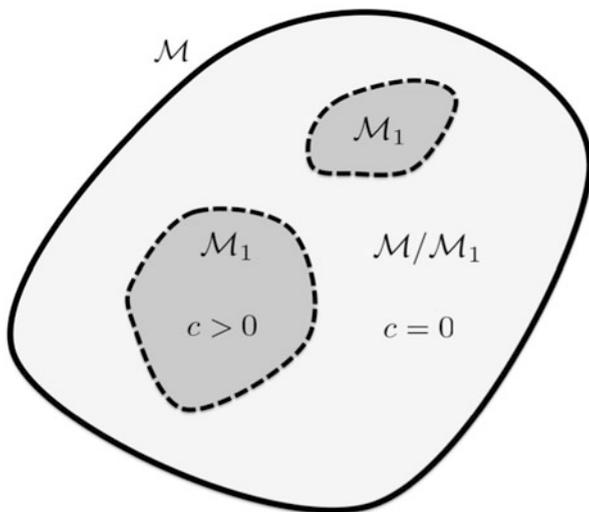
Theorem 3 *The statistical power \mathcal{P} of testing the hypotheses*

$$H_0 : h(p) = 0 \text{ for all } p \in \mathcal{M} \text{ vs. } H_1 : h(p) = c\sigma > 0 \text{ for some } p \in \mathcal{M}.$$

using the T random field $T(p)$ is given by

$$\mathcal{P}(n) \approx 1 - \exp \left[- \sum_{j=0}^d \mu_j(\mathcal{M}_1) \rho_j(t_\alpha^* - c\sqrt{n}) \right],$$

Fig. 6 Schematic of the hypothesis testing in Theorem 3 when H_1 is true. Since the hyoid bone is composed of three structures (hyoid body and two greater horns), we can have multiple disconnected \mathcal{M}_1 , where $c > 0$



where t_α^* is the α -quantile given by

$$\alpha = P\left(\sup_{p \in \mathcal{M}} T(p) > t_\alpha^*\right).$$

Proof In the region $\mathcal{M}_0 = \mathcal{M} \setminus \mathcal{M}_1$ corresponding to H_0 ,

$$f^i(p) \sim N(0, \sigma^2).$$

In the region \mathcal{M}_1 corresponding to H_1 ,

$$f^i(p) \sim N(c\sigma, \sigma^2).$$

Figure 6 illustrates this setting, where \mathcal{M}_1 can be disconnected sets. Consider the test statistic

$$T(p) = \frac{\bar{f}(p)}{S(p)/\sqrt{n}}, \tag{20}$$

where \bar{f} and S are the sample mean and standard deviation of the measurements f^1, \dots, f^n . In \mathcal{M}_0 , $T(p)$ is a T random field with $n - 1$ degrees of freedom [1]. In \mathcal{M}_1 , $T(p)$ can be written as

$$T(p) = T'(p) + \frac{c\sigma}{S(p)/\sqrt{n}},$$

where $T'(p)$ a T random field with $n - 1$ degrees of freedom. Since σ is usually estimated using the standard deviation, approximately, we have $S(p) = \sigma$ and the test statistic becomes $T(p) = T'(p) + c\sqrt{n}$ in \mathcal{M}_1 . At each fixed p , $T(p)$ is no longer a T random field but a non-central T random field [30]. Subsequently the power \mathcal{P} at the α -level is given by

$$\mathcal{P}(n) = P\left(\sup_{p \in \mathcal{M}_1} T(p) > t_\alpha^*\right) \quad (21)$$

$$= P\left(\sup_{p \in \mathcal{M}_1} T'(p) > t_\alpha^* - c\sqrt{n}\right), \quad (22)$$

where t_α^* is the α -quantile of $\sup_{p \in \mathcal{M}} T(p)$ under H_0 , i.e.

$$\alpha = P\left(\sup_{p \in \mathcal{M}} T(p) > t_\alpha^*\right).$$

Although (22) is intractable to directly compute, we can approximate (22) using the expected Euler characteristic (EC) [59, 61]. The power (22) can be written as

$$\mathcal{P}(n) = \sum_{j=0}^d \mu_j(\mathcal{M}_1) \rho_j(t_\alpha^* - c\sqrt{n}),$$

where $\mu_d(\mathcal{M})$ is the j -th Minkowski functional or intrinsic volume of \mathcal{M} and ρ_j is the j -th EC-density of T-field [1, 54, 59, 60]. The expansion only works for sufficiently large $t_\alpha^* - c\sqrt{n}$. The rate of the convergence is given in terms of probability as $O((t_\alpha^*)^{-1/2})$ [58]. For small thresholds, the power may not be bounded between 0 and 1. Thus, it is necessary to use the exponential transform to bound the power [30]. For small $\mathcal{P}(n)$, using the Taylor expansion, we can write $\exp[-\mathcal{P}(n)] \approx 1 - \mathcal{P}(n)$. Equivalently, it is written as $\mathcal{P}(n) \approx 1 - \exp[-\mathcal{P}(n)]$. This transformation guarantees the power estimation to be bounded between 0 and 1 [30]. Subsequently, the power is given by

$$\mathcal{P}(n) = 1 - \exp\left[-\sum_{j=0}^d \mu_j(\mathcal{M}_1) \rho_j(t_\alpha^* - c\sqrt{n})\right]. \quad (23)$$

Figure 7 displays the power $\mathcal{P}(n)$ over the sample size n for effect sizes $c = 0.1, 0.2, 0.5$ based on (23). The actual surface of the hyoid bone is taken as \mathcal{M} and 10% of surface area is taken as the signal region \mathcal{M}_1 .

□

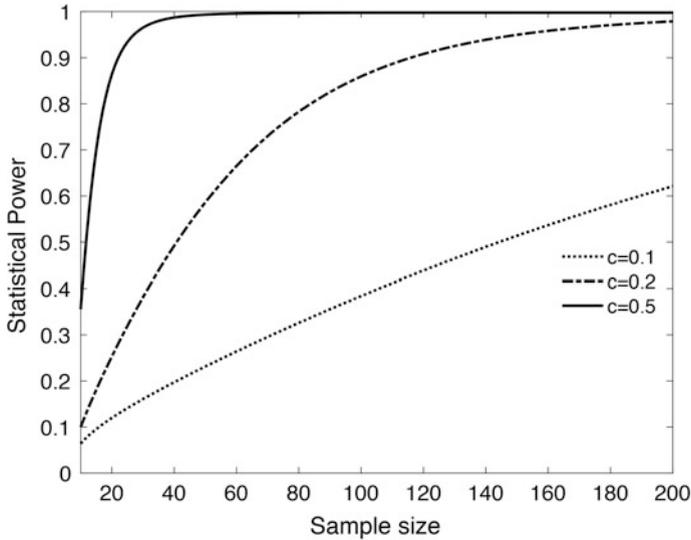


Fig. 7 Statistical power over the sample size under the multiple comparisons. c is the effect size using formula (23). The surface of the hyoid bone is taken as \mathcal{M} and 10% of surface area is taken as the signal region \mathcal{M}_1

3.5 Validation

The proposed method is validated against the *iterated kernel smoothing* [17, 18], which smooth data as weighted average of neighboring mesh vertices using a Gaussian kernel and its iterative application is supposed to approximate the diffusion process. The iterated Gaussian kernel smoothing was also used as the baseline method in [40]. We performed two simulations with small and large signal to noise ratio (SNR) settings on a T-junction surface with three different curvatures: convex, concave and almost flat regions (Fig. 8). Surface smoothing methods perform differently under different curvatures. Three signal regions of different sizes (colored red in Fig. 8) were taken as the ground truth at these regions and 60 independent functional measurements on the surface were simulated as $|N(0, \gamma^2)|$, the absolute value of normal distribution with mean 0 and variance γ^2 , at each mesh vertex. Value 1 was then added to the regions in 30 of the measurements, which served as group II, while the other 30 measurements were taken as group I. Group I has distribution $|N(0, \gamma^2)|$ while group II has distribution $|N(1, \gamma^2)|$ in the signal regions. Larger variance γ^2 corresponds to smaller SNR.

In Study I, $\gamma^2 = 2^2$ was used to simulate smaller SNR. Figure 8 shows the simulation results. For iterated kernel smoothing [17, 18], we used bandwidth $t = 0.5$ and 100 iterations (second column). The expansion with 1000 LB eigenfunction is used to smooth data, which is equivalent to heat kernel smoothing with zero bandwidth (third column). For heat kernel smoothing, bandwidth $t = 0.5$ and 1000

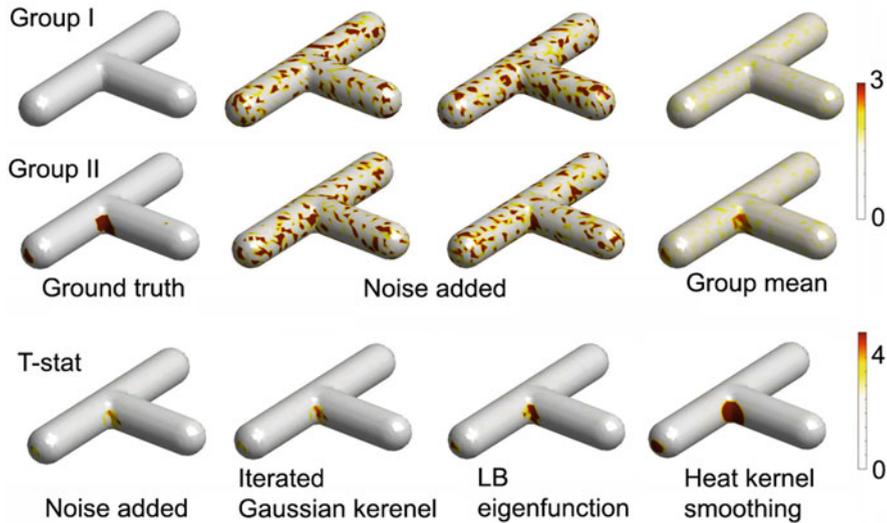


Fig. 8 Simulation study on a T-junction shaped surface where three regions of different sizes are taken as the ground truth (colored red in group II - ground truth). 60 independent functional measurements on the T-junction were simulated as $|N(0, \gamma^2)|$ at each mesh vertex. Value 1 was added to the ground truth region in 30 measurements, which served as group II while the other 30 measurements were taken as group I. T-statistics are shown for these simulations (original) and three techniques with bandwidth 0.5. Heat kernel smoothing performed the best in detecting the ground truth regions

eigenfunctions were used (fourth column). We then performed a two sample t -test with the random field theory corrected threshold of 4.90 to detect the group difference at $\alpha = 0.05$ level. The noise added raw data were able to correctly identify only 3% of signal regions but also detected 3% of non-signal regions as signal. Iterated kernel smoothing also was able to identify only 3% of signal regions as signal but also detected 3% of non-signal regions as signal. The LB eigenfunction expansion were able to correctly identify 25% of signal regions but did not detect any signal in non-signal regions as signal. In comparison, heat kernel correctly identified 94% of the signal regions and incorrectly identified 0.4% of non-signal regions as signal. The proposed heat kernel smoothing performed very well in the small SNR setting.

In Study II, $\gamma^2 = 1$ was used to simulate functional measurements with substantially larger SNR. The same parameters were used as in Study I. The noise added raw data was able to correctly identify 88% of signal regions and did not detect any signal in non-signal regions as signal. Iterated kernel smoothing was able to correctly identify 91% of signal regions and did not detect any signal in non-signal regions as signal. LB eigenfunction expansion was able to correctly identify only 94% of signal regions and did not detect any signal in non-signal regions as signal. In comparison, heat kernel correctly identified 97% of the signal regions and

incorrectly identified 1.5% of non-signal regions as signal. Although all the methods performed well in small SNR setting, the proposed heat kernel smoothing performed the best.

4 Application

4.1 CT Imaging Data and Preprocessing

This study consists of high resolution CT images of 70 typically developing individuals ages between 0 and 20 years (mean age 8.0 ± 11.3 years). CT scans were converted to DICOM format and Analyze 8.1 software package (AnalyzeDirect, Inc., Overland Park, KS) was then used in segmenting binary hyoid bone images by a trained individual rater in the native space by simple image intensity thresholding and careful manual editing [10, 16]. A nonlinear image registration using the diffeomorphic shape and intensity averaging technique with cross-correlation as similarity metric was performed through Advanced Normalization Tools (ANTs) [4]. Some individual may have larger hyoid than others so it was necessary to remove the global size differences in local shape modeling. From the affine transformed individual hyoid surfaces, we performed the diffeomorphic nonlinear image registration to the template. A study-specific template was constructed as follows. We chose a 12 year old female identified as F155 as the initial template and aligned the remaining 69 hyoids to this template affinely to remove the overall size variability. F155 was carefully chosen among all other segmentation results by visual inspection to have no segmentation artifacts. Further, it was constantly used as a reference template in previous studies [49, 50]. By averaging the inverse deformation fields from the initial template to individual hyoid, we obtained the yet another final template. Since the final template is the average of all other surfaces, the final localized growth pattern is not much influenced by the choice of the initial template.

Image acquisition error, discretization error, and image preprocessing noises in segmentation and registration often result in noisy deformation fields. The proposed heat kernel smoothing was applied to the displacement vector fields to smooth out high frequency noises. 70 individuals are binned into three age groups: 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, 14 and 30 individuals in group I, II and III respectively. The main biological hypothesis of interest is if there is any localized hyoid bone growth spurts between these specific age groups. The age range is chosen based on prior bone growth studies [35], where similar age binning is used in modeling the growth of mandible, which is located in the close proximity to the hyoid bone.

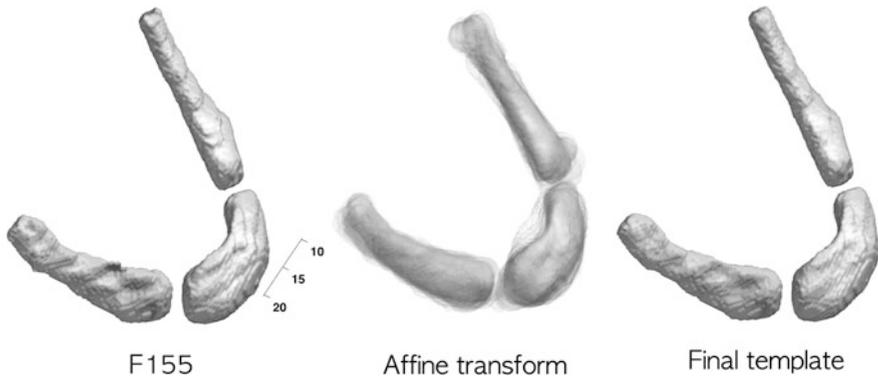


Fig. 9 Left: Hyoid F155 which forms an initial template \mathcal{M}_I . All other hyoids are affine registered to F155. Middle: The superimposition of affine registered hyoids showing local misalignments. Diffeomorphic registration is then performed to register misaligned affine transformed hyoids. Right: The average of deformation with respect to F155 provides the final population average template \mathcal{M}_F where statistical parametric maps will be constructed

4.2 Results

Figure 9 shows the initial and final templates. The isosurface of the final template volume is extracted using the marching cubes algorithm [41]. The displacement from the template to an individual surface is obtained at each mesh vertex. Figure 10 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. We are interested in localizing the regions of hyoid bone growth between the age groups.

Since the length measurement provides a much easier biological interpretation, we used the length of displacement vector as a response variable among many other possible features. Since the length on the template surface is expected to be noisy due to image acquisition, segmentation and image registration errors, it is necessary to perform the proposed kernel regression and subsequently reduce the type-I error and obtain more stable SPM. Figure 3 shows an example of kernel smoothing on our data. The kernel smoothing increases the signal-to-noise ratio (SNR) and improves the smoothness and Gaussianness of data. Subsequently, the heat kernel smoothing of the displacement length is taken as the response variable. We have chosen $t = 5$ as the bandwidth for the study since the bandwidth 5 is where the type-I error starts to flatten out in Fig. 5. Note that the LB-eigenfunction expansion with 500 eigenfunctions is close to the original data (relative error of less than 0.3%). Hence, performing the proposed kernel regression before the statistical analysis can substantially smaller type-I error demonstrating its effectiveness.



Fig. 10 Hyoid bones 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) and the mean displacements between the groups are visualized. 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

After the displacement lengths are smoothed, we constructed the F-field, or equivalently the T-field square, for testing the length difference between the age groups I and II, II and III, and I and III showing the regions of growth spurts between different age range (Fig. 11). Since test statistics are constructed over all mesh vertices on the hyoid bone, multiple comparisons were account for using the random field theory [62, 63].

For testing the differences between the groups I and II, II and III, and I and III, they are based on F-field with 1 and 38, 1 and 42, and 1 and 54 degrees of freedom respectively. The result is displayed in Fig. 11, where the significant results were only found between the groups II and III (middle), and I and III (bottom) at $\alpha = 0.1$ level. Between the groups I and II, we obtained maximum F -statistic value of 4.58 (left hyoid), which is not significant enough. Between the groups II and III, we the maximum F -statistic value of 9.36 (right hyoid), which corresponds to the p -value of 0.13 (corrected). Between the groups I and III, we obtained the maximum F -statistic value of 10.55 (middle hyoid), which corresponds to the p -value of 0.074 (corrected). The multiple comparisons were done over the whole hyoid bone. If we perform the multiple comparisons for each of the three components of the hyoid bone, we can boost the signal a bit. For instance, restricted to the middle hyoid bone, the maximum F -statistic value of 10.55 will correspond to the p -value of 0.028.

F-stat (Group II - I)



F-stat (Group III - II)



F-stat (Group III - I)



Fig. 11 F-statistic maps on hyoid showing age effect between the groups. The significant growth regions (red) are identified only between groups II and III, and I and III. The growth is highly localized near the regions that connect the disconnected hyoid bones

5 Conclusions

We have developed a new kernel regression framework on a manifold that unifies bivariate kernel regression, heat diffusion and wavelets in a single coherent mathematical framework. The kernel regression is robust both globally and locally in that it uses global basis functions to perform regression but locally related to the diffusion wavelet transform. The proposed framework is demonstrated to reduce the type-I error in modeling shape variations compared to the usual LB-eigenfunction expansion. The method is then used in developing a statistical inference procedure for functional signals on manifolds.

Acknowledgements This work was supported by NIH Research Grants R01 DC6282, R01 EB022856, UL1TR000427 and P-30 HD03352 and U54 HD090256 to the Waisman Center. We thank Vikas Singh and Won Hwa Kim of University of Wisconsin-Madison for discussion on wavelets.

References

1. Adler, R. J. (1981). *The geometry of random fields*. London: Wiley.
2. Andrade, A., Kherif, F., Mangin, J., Worsley, K. J., Paradis, A., Simon, O., et al. (2001). Detection of fMRI activation using cortical surface mapping. *Human Brain Mapping, 12*, 79–93.
3. Antoine, J.-P., Roşca, D., & Vandergheynst, P. (2010). Wavelet transform on manifolds: Old and new approaches. *Applied and Computational Harmonic Analysis, 28*, 189–202.
4. Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis, 12*, 26–41.
5. Belkin, M., & Niyogi, P. (2002). Laplacian eigenmaps and spectral techniques for embedding and clustering. In *Advances in Neural Information Processing Systems* (pp. 585–592)
6. Belkin, M., Niyogi, P., & Sindhvani, V. (2006). Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *The Journal of Machine Learning Research, 7*, 2399–2434.
7. Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics, 29*, 1165–1188.
8. 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.
9. Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., & Brammer, M. J. (1999). Global, voxel, and cluster tests, by theory and permutation, for adifference between two groups of structural MR images of the brain. *IEEE Transactions on Medical Imaging, 18*, 32–42.
10. Chuang, Y. J., Doherty, B. M., Adluru, N., Chung, M. K., & Vorperian, H. K. (2018). A novel registration-based semiautomatic mandible segmentation pipeline using computed tomography images to study mandibular development. *Journal of Computer Assisted Tomography, 42*, 306–316.
11. Chung, M.K. (2012). *Computational neuroanatomy: The methods*. Singapore: World Scientific.
12. Chung, M. K., Dalton, K.M., & Davidson, R. J. (2008). Tensor-based cortical surface morphometry via weighted spherical harmonic representation. *IEEE Transactions on Medical Imaging, 27*, 1143–1151.
13. Chung, M. K., Dalton, K. M., Shen, L., Evans, A. C., & Davidson, R. J. (2007). Weighted Fourier representation and its application to quantifying the amount of gray matter. *IEEE Transactions on Medical Imaging, 26*, 566–581.
14. Chung, M. K., Hartley, R., Dalton, K. M., & Davidson, R. J. (2008). Encoding cortical surface by spherical harmonics. *Statistica Sinica, 18*, 1269–1291.
15. Chung, M. K., Lee, H., Solo, V., Davidson, R. J., & Pollak, S. D. (2017). Topological distances between brain networks. *International Workshop on Connectomics in Neuroimaging, Lecture Notes in Computer Science* (pp. 161–170).
16. Chung, M. K., Qiu, A., Seo, S., & Vorperian, H. K. (2015). Unified heat kernel regression for diffusion, kernel smoothing and wavelets on manifolds and its application to mandible growth modeling in CT images. *Medical Image Analysis, 22*, 63–76.

17. Chung, M. K., Robbins, S., Dalton, K. M., Davidson, R. J., Alexander, A. L., & Evans, A. C. (2005). Cortical thickness analysis in autism with heat kernel smoothing. *NeuroImage*, *25*, 1256–1265.
18. Chung, M. K., Robbins, S., & Evans, A. C. (2005). Unified statistical approach to cortical thickness analysis. In *Information Processing in Medical Imaging (IPMI). Lecture Notes in Computer Science* (Vol. 3565, pp. 627–638)
19. Chung, M. K., & Taylor, J. (2004). Diffusion smoothing on brain surface via finite element method. In *Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI)* (Vol. 1, pp. 432–435).
20. Chung, M. K., Vilalta-Gil, V., Lee, H., Rathouz, P. J., Lahey, B. B., & Zald, D. H. (2017). Exact topological inference for paired brain networks via persistent homology. In *Information Processing in Medical Imaging (IPMI). Lecture Notes in Computer Science* (Vol. 10265, pp. 299–310).
21. Chung, M. K., Worsley, K. J., Paus, T., Cherif, D. L., Collins, C., Giedd, J., et al. (2001). A unified statistical approach to deformation-based morphometry. *NeuroImage*, *14*, 595–606.
22. Chung, M. K., Worsley, K. J., Robbins, S., Paus, T., Taylor, J., Giedd, J. N., et al. (2003). Deformation-based surface morphometry applied to gray matter deformation. *NeuroImage*, *18*, 198–213.
23. Courant, R., & Hilbert, D. (1953). *Methods of mathematical physics*. New York: Interscience. English edition.
24. Fan, J., & Gijbels, I. (1996). *Local polynomial modelling and its applications*. Boca Raton: Chapman & Hall/CRC.
25. Fletcher, T. (2011). Geodesic regression on Riemannian manifolds. In *Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy-Geometrical and Statistical Methods for Modelling Biological Shape Variability* (pp. 75–86).
26. Gelb, A. (1997). The resolution of the Gibbs phenomenon for spherical harmonics. *Mathematics of Computation*, *66*, 699–717.
27. Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*, 870–878.
28. Gu, X., Wang, Y. L., Chan, T. F., Thompson, T. M., & Yau, S. T. (2004). Genus zero surface conformal mapping and its application to brain surface mapping. *IEEE Transactions on Medical Imaging*, *23*, 1–10.
29. Hammond, D. K., Vandergheynst, P., & Gribonval, R. (2011). Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, *30*, 129–150.
30. Hayasaka, S., Peiffer, A. M., Hugenschmidt, C. E., & Laurienti, P. J. (2007). Power and sample size calculation for neuroimaging studies by non-central random field theory. *NeuroImage*, *37*, 721–730.
31. Hayasaka, S., Phan, K. L., Liberzon, I., Worsley, K. J., & Nichols, T. E. (2004). Nonstationary cluster-size inference with random field and permutation methods. *Neuroimage*, *22*, 676–687.
32. Hendriks, H. (1990). Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansions. *The Annals of Statistics*, *18*, 832–849.
33. Hosseinbor, A. P., Kim, W. H., Adluru, N., Acharya, A., Vorperian, H. K., & Chung, M. K. (2014). The 4D hyperspherical diffusion wavelet: A new method for the detection of localized anatomical variation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Vol. 8675, pp. 65–72).
34. Hurdal, M. K., & Stephenson, K. (2004). Cortical cartography using the discrete conformal approach of circle packings. *NeuroImage*, *23*, S119–S128.
35. Kelly, M. P., Vorperian, H. K., Wang, Y., Tillman, K. K., Werner, H. M., Chung, M. K., et al. (2017). Characterizing mandibular growth using three-dimensional imaging techniques and anatomic landmarks. *Archives of Oral Biology*, *77*, 27–38.
36. Kim, H. J., Adluru, N., Collins, M. D., Chung, M. K., Bendlin, B. B., Johnson, S. C., et al. (2014). Multivariate general linear models (MGLM) on Riemannian manifolds with applications to statistical analysis of diffusion weighted images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2705–2712).

37. Kim, S.-G., Chung, M. K., Seo, S., Schaefer, S. M., van Reekum, C., & Davidson, R. J. (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* (Vol. 7087, pp. 36–47).
38. Kim, W. H., Pachauri, D., Hatt, C., Chung, M. K., 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).
39. Lerch, J. P., & Evans, A. C. (2005). Cortical thickness analysis examined through power analysis and a population simulation. *NeuroImage*, *24*, 163–173.
40. Lila, E., Aston, J. A. D., & Sangalli, L. M. (2016). Smooth principal component analysis over two-dimensional manifolds with an application to neuroimaging. *The Annals of Applied Statistics*, *10*, 1854–1879.
41. Lorensen, W. E., & Cline, H. E. (1987). Marching cubes: A high resolution 3D surface construction algorithm. In *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques* (pp. 163–169).
42. Mahadevan, S., & Maggioni, M. (2006). Value function approximation with diffusion wavelets and laplacian eigenfunctions. *Advances in Neural Information Processing Systems*, *18*, 843.
43. Malladi, R., & Ravve, I. (2002). Fast difference schemes for edge enhancing Beltrami flow. In *Proceedings of Computer Vision-ECCV. Lecture Notes in Computer Science* (Vol. 2350, pp. 343–357).
44. Nain, D., Styner, M., Niethammer, M., Levitt, J. J., Shenton, M. E., G. Gerig, et al. (2007). Statistical shape analysis of brain structures using spherical wavelets. In *IEEE Symposium on Biomedical Imaging ISBI* (Vol. 4, pp. 209–212).
45. Öztireli, A. C., Guennebaud, G., & Gross, M. (2009). Feature preserving point set surfaces based on non-linear kernel regression. In *Computer Graphics Forum* (Vol. 28, pp. 493–501).
46. Perona, P., & Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence*, *12*, 629–639.
47. Qiu, A., Bitouk, D., & Miller, M.I. (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.
48. Qiu, A., & Miller, M. I. (2008). Multi-structure network shape analysis via normal surface momentum maps. *NeuroImage*, *42*, 1430–1438.
49. Seo, S., Chung, M. K., & Vorperian, H. K. (2010). Heat kernel smoothing using Laplace-Beltrami eigenfunctions. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2010. Lecture Notes in Computer Science* (Vol. 6363, pp 505–512).
50. Seo, S., Chung, M. K., & Voperian, H. K. (2011). Mandible shape modeling using the second eigenfunction of the Laplace-Beltrami operator. In *Proceedings of SPIE Medical Imaging* (Vol. 7962, pp. 79620Z).
51. Sochen, N., Kimmel, R., & Malladi, R. (1998). A general framework for low level vision. *IEEE Transactions on Image Processing*, *7*, 310–318.
52. Tang, B., Sapiro, G., & Caselles, V. (1999). Direction diffusion. In *The Proceedings of the Seventh IEEE International Conference on Computer Vision* (Vol. 2, pp. 1245–1252).
53. Taubin, G. (2000). Geometric signal processing on polygonal meshes. In *EUROGRAPHICS*. Geneva: Eurographics Association.
54. Taylor, J. E., & Worsley, K. J. (2007). Detecting sparse signals in random fields, with an application to brain mapping. *Journal of the American Statistical Association*, *102*, 913–928. .
55. Vorperian, H. K., Wang, S., Schimek, E. M., Durtschi, R. B., Kent, R. D., Gentry, L. R., et al. (2011). Developmental sexual dimorphism of the oral and pharyngeal portions of the vocal tract: an imaging study. *Journal of Speech, Language and Hearing Research*, *54*, 995–1010. .
56. Wang, Y., Chung, M. K., & Vorperian, H. K. (2016). Composite growth model applied to human oral and pharyngeal structures and identifying the contribution of growth types. *Statistical Methods in Medical Research*, *25*, 1975–1990.

57. Wang, Y., Zhang, J., Gutman, B., Chan, T. F., Becker, J. T., Aizenstein, H. J., et al. (2010). Multivariate tensor-based morphometry on surfaces: Application to mapping ventricular abnormalities in HIV/AIDS. *NeuroImage*, *49*, 2141–2157.
58. Worsley, K. J. (1994). Local maxima and the expected Euler characteristic of excursion sets of χ^2 , f and t fields. *Advances in Applied Probability*, *26*, 13–42.
59. Worsley, K. J. (2003). Detecting activation in fMRI data. *Statistical Methods in Medical Research*, *12*, 401–418.
60. Worsley, K. J., Cao, J., Paus, T., Petrides, M., & Evans, A. C. (1998). Applications of random field theory to functional connectivity. *Human Brain Mapping*, *6*, 364–367.
61. Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, *4*, 58–73.
62. Worsley, K. J., Poline, J.-B., Vandal, A. C., & Friston, K. J. (1995). Test for distributed, non-focal brain activations. *NeuroImage*, *2*, 173–181.
63. Worsley, K. J., Taylor, J. E., Tomaiuolo, F., & Lerch, J. (2004). Unified univariate and multivariate random field theory. *NeuroImage*, *23*, S189–195.
64. Xu, Y., Valentino, D. J., Scher, A. I., Dinov, I., White, L. R., Thompson, P. M., Launer, L. J., & Toga, A. W. (2008). Age effects on hippocampal structural changes in old men: The HAAS. *NeuroImage*, *40*, 1003–1015.
65. Yushkevich, P. A., Zhang, H., Simon, T. J., & Gee, J. C. (2008). Structure-specific statistical mapping of white matter tracts. *NeuroImage*, *41*, 448–461.
66. Zhang, H., Avants, B. B., Yushkevich, P. A., Woo, J. H., Wang, S., McCluskey, L. F., et al. (2007). High-dimensional spatial normalization of diffusion tensor images improves the detection of white matter differences: an example study using amyotrophic lateral sclerosis. *IEEE Transactions on Medical Imaging*, *26*(11), 1585–1597.