

Weighted Fourier Series Representation and Its Application to Quantifying the Amount of Gray Matter

Moo K. Chung*, Kim M. Dalton, Li Shen, Alan C. Evans, and Richard J. Davidson

Abstract—We present a novel weighted Fourier series (WFS) representation for cortical surfaces. The WFS representation is a data smoothing technique that provides the explicit smooth functional estimation of unknown cortical boundary as a linear combination of basis functions. The basic properties of the representation are investigated in connection with a self-adjoint partial differential equation and the traditional spherical harmonic (SPHARM) representation. To reduce steep computational requirements, a new iterative residual fitting (IRF) algorithm is developed. Its computational and numerical implementation issues are discussed in detail. The computer codes are also available at <http://www.stat.wisc.edu/~mchung/softwares/weighted-SPHARM/weighted-SPHARM.html>. As an illustration, the WFS is applied in quantifying the amount of gray matter in a group of high functioning autistic subjects. Within the WFS framework, cortical thickness and gray matter density are computed and compared.

Index Terms—Cortical thickness, diffusion smoothing, gray matter density, iterative residual fitting, SPHARM, spherical harmonics.

I. INTRODUCTION

IN this paper, we present a new morphometric framework called the *weighed Fourier series* (WFS) representation. The WFS is both a global hierarchical parameterization and an explicit data smoothing technique formulated as a solution to a self-adjoint partial differential equation (PDE). WFS generalizes the traditional spherical harmonic (SPHARM) representation [23], [44] with additional exponential weights. The exponentially decaying weights make the representation converges faster and reduce the ringing artifacts (Gibbs phenomenon) [22] significantly. Unlike SPHARM, WFS can be reformulated as heat kernel smoothing [9] when the self-adjoint operator becomes the Laplace–Beltrami operator. Then as similar to heat kernel smoothing, the random field theory

[51]–[53] can be used for statistical inference on localizing abnormal shape variations in a clinical population. Many basic theoretical properties of WFS and its numerical implementation issues are presented in great detail. The WFS representation requires estimating 18 723 unknown Fourier coefficients on a high resolution spherical mesh sampled at more than 40 000 vertices. This requires a specialized linear solver with fairly steep computational resources. To address this issue, we have developed a new estimation technique called the *iterative residual fitting* (IRF) algorithm [43]. The computational burden has been reduced substantially by decomposing the subspace spanned by spherical harmonics into smaller subspaces, and iteratively performing the least squares estimation on the smaller subspaces. The correctness of the algorithm is proven and its accuracy is numerically evaluated.

As an illustration of the new technique, we have applied it in localizing abnormal amounts of gray matter in a group of high functioning autistic subjects. The cerebral cortex has a highly convoluted geometry and it is likely that the local difference in gray matter concentration can characterize a clinical population. Within the WFS framework, gray matter density and cortical thickness are also computed. Then statistical parametric maps (SPM) are constructed to localize the regions of abnormal gray matter. These two measurements are also compared to determine if increased cortical thickness corresponds to increased gray matter locally.

The main contributions of the paper are the development of the underlying theory of the WFS representation and its numerical implementation using the iterative residual fitting (IRF) algorithm. We also made the computer code freely available to public. In the following subsections, we will review the literature that is directly related to our methodology and address what our specific contributions are in the context of the previous literature.

A. Spherical Harmonic Representation

The SPHARM representation [4] has been applied to subcortical structures such as the hippocampus and the amygdala [23], [27], [30], [44]. In particular, Gerig *et al.* used the mean squared distance (MSD) of the SPHARM coefficients in quantifying ventricle surface shape in a twin study [23]. Shen *et al.* used the principal component analysis technique on the SPHARM coefficients of schizophrenic hippocampal surfaces in reducing the data dimension [44]. Recently, it has begun to be applied to more complex cortical surfaces [27], [43]. Gu *et al.* presented SPHARM as a surface compression technique, where the main

Manuscript received October 1, 2006; revised December 29, 2006. Asterisk indicates corresponding author.

*M. K. Chung is with the Department of Statistics, Biostatistics, and Medical Informatics, and the Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin, Madison, WI 53706 USA (e-mail: mchung@stat.wisc.edu).

K. M. Dalton and R. J. Davidson are with the Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin, Madison, WI 53706 USA.

L. Shen is with the Department of Computer and Information Science, University of Massachusetts, Dartmouth, MA 02747 USA.

A. C. Evans is with the McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, QC H3A 2B4 Canada.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2007.892519

geometric features are encoded in the low degree spherical harmonics, while the noises are in the high degree spherical harmonics [27].

In SPHARM, the spherical harmonic are used in constructing the Fourier series expansion of the mapping from cortical surfaces to a unit sphere. So SPHARM is more of an interpolation technique than a smoothing technique, and thus it will have the ringing artifacts [22]. On the other hand, WFS is a kernel smoothing technique given as a solution to a self-adjoint PDE. The solution to the PDE is expanded in basis functions. In a similar spirit, Bulow used the spherical harmonics in isotropic heat diffusion via the Fourier transform on a unit sphere as a form of hierarchical surface representation [5]. WFS offers many advantages over the previous PDE-based smoothing techniques [1], [10]. The PDE-based smoothing methods tend to suffer numerical instability [1], [6], [10] while WFS has no such problem. Since the traditional PDE-based smoothing gives an implicit numerical solution, setting up a statistical model is not straightforward. However, WFS provides an explicit series expansion so it is easy to apply a wide variety of statistical modeling techniques such as the GLM [21], principal component analysis (PCA) [44] and functional-PCA [36], [41].

SPHARM will be shown to be the special case of WFS. In the SPHARM representation, all measurements are assigned equal weights and the coefficients of the series expansion is estimated in the least squares fashion. In WFS, closer measurements are weighted more and the coefficients of the series expansion is estimated in the weighted least squares fashion. So WFS is more suitable than SPHARM when the realization of the cortical boundaries, as triangle meshes, are noisy and possibly discontinuous. In most SPHARM literature, the degree of the Fourier series expansion has been arbitrarily determined and the problem of the optimal degree has not been addressed. Our WFS formulation addresses the determination of the optimal degree in a unified statistical modeling framework. The WFS-based global parametrization is computationally expensive compared to the local quadratic polynomial fitting [4], [10], [13], [29], [40] while providing more accuracy and flexibility for hierarchical representation.

B. Cortical Thickness

The cerebral cortex has the topology of a 2-D convoluted sheet. Most of the features that distinguish these cortical regions can only be measured relative to that local orientation of the cortical surface [13]. The CSF and gray matter interface is referred to as the *outer surface* (pial surface) while the gray and white matter interface is referred to as the *inner surface* [10], [32]. Then the distance between the outer and inner surfaces is defined as the *cortical thickness* and it has been widely used as an anatomical index for quantifying the amount of gray matter in the brain [9], [10], [18].

Unlike 3-D whole brain volume based gray matter density, 1-D cortical thickness measures have the advantage of providing a direct quantification of cortical geometry. It is likely that different clinical populations will exhibit different cortical thickness. By analyzing cortical thickness, brain shape differences can be quantified locally [10], [19], [32], [35]. The cortical surfaces are usually segmented as triangle meshes that are con-

structed from deformable surface algorithms [15], [18], [32]. Then the cortical thickness is mainly defined and estimated as the shortest distance between vertices of the two triangle meshes [18], [32]. The mesh construction and discrete thickness computation procedures introduce substantial noise in the thickness measure [9] (Fig. 6). So it is necessary to increase the signal-to-noise ratio (SNR) and smoothness of data for the subsequent random field based statistical analysis. For smoothing cortical data, diffusion equation based methods have been used [1], [6], [9], [10]. The shortcoming of these approaches is the need for numerically solving the diffusion equation possibly via the finite element technique. This is an additional computational step on top of the cortical thickness estimation. In this paper, we present a more direct approach that smoothes and parameterizes the coordinates of a mesh directly via WFS such that the resulting thickness measures are already smooth. In WFS, the cortical surfaces are estimated as a weighted linear combination of smooth basis functions so that most algebraic operations on WFS will also be smooth.

II. CAUCHY PROBLEM AS A SMOOTHING PROCESS

Consider $\mathcal{M} \in \mathbb{R}^d$ to be a compact differentiable manifold. Let $L^2(\mathcal{M})$ be the space of square integrable functions in \mathcal{M} with inner product

$$\langle g_1, g_2 \rangle = \int_{\mathcal{M}} g_1(p) g_2(p) d\mu(p) \quad (1)$$

where μ is the Lebesgue measure such that $\mu(\mathcal{M})$ is the total volume of \mathcal{M} . The norm $\| \cdot \|$ is defined as

$$\|g\| = \langle g, g \rangle^{1/2}.$$

The linear partial differential operator \mathcal{L} is *self-adjoint* if

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

for all $g_1, g_2 \in L^2(\mathcal{M})$. Then the eigenvalues λ_j and eigenfunctions ψ_j of the operator \mathcal{L} are obtained by solving

$$\mathcal{L}\psi_j = \lambda_j \psi_j. \quad (2)$$

Without the loss of generality, we can order eigenvalues

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

and the eigenfunctions to be orthonormal with respect to the inner product (1). Consider a Cauchy problem of the following form:

$$\partial_t g + \mathcal{L}g = 0, \quad g(p, t = 0) = f(p). \quad (3)$$

The initial functional data $f(p)$ can be further stochastically modeled as

$$f(p) = \nu(p) + \epsilon(p) \quad (4)$$

where ϵ is a stochastic noise modeled as a mean zero Gaussian random field and ν is the unknown signal to be estimated. The

PDE (3) diffuses noisy initial data f over time and estimate the unknown signal ν as a solution. The time t controls the amount of smoothing and will be termed as the *bandwidth*. The unique solution to (3) is given by the following theorem.

Theorem 1: For the self-adjoint linear differential operator \mathcal{L} , the unique solution of the Cauchy problem (3) is given by

$$g(p, t) = \sum_{j=0}^{\infty} e^{-\lambda_j t} \langle f, \psi_j \rangle \psi_j(p). \quad (5)$$

Proof: For each fixed t , g has expansion

$$g(p, t) = \sum_{j=0}^{\infty} c_j(t) \psi_j(p). \quad (6)$$

Substitute (6) into (3). Then we obtain

$$\partial_t c_j(t) + \lambda_j c_j(t) = 0. \quad (7)$$

The solution of (7) is given by $c_j(t) = b_j e^{-\lambda_j t}$. So we have 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 $\langle f, \psi_j \rangle$.

The implication of Theorem 1 is obvious. The solution decreases exponentially as time t increases and smoothes out high spatial frequency noise much faster than low-frequency noise. This is the basis of many of PDE-based image smoothing methods. PDE involving self-adjoint linear partial differential operators such as the Laplace–Beltrami operator or iterated Laplacian have been widely used in medical image analysis as a way to smooth either scalar or vector data along anatomical boundaries [1], [5], [6], [10]. These methods directly solve the PDE using standard numerical techniques such as the finite difference method or the finite element method. The problem with directly solving PDEs is the numerical instability and the complexity of setting up the numerical scheme. WFS differs from these previous methods in such a way that we only need to estimate the Fourier coefficients in a hierarchical fashion to solve the PDE.

A. Weighted Fourier Series

We will investigate the properties of the finite expansion of (5) denoted by

$$\mathcal{F}_t^k[f](p) = \sum_{j=0}^k e^{-\lambda_j t} \langle f, \psi_j \rangle \psi_j(p).$$

This expansion will be called as the *weighted Fourier Series* (WFS). By rearranging the inner product, the WFS can be rewritten as kernel smoothing

$$\mathcal{F}_t^k[f](p) = \sum_{j=0}^k e^{-\lambda_j t} \psi_j(p) \int_{S^2} f(q) \psi_j(q) d\mu(q) \quad (8)$$

$$= \int_{S^2} f(q) K_t^k(p, q) d\mu(q) \quad (9)$$

with symmetric positive definite kernel K_t^k given by

$$K_t^k(p, q) = \sum_{j=0}^k e^{-\lambda_j t} \psi_j(p) \psi_j(q). \quad (10)$$

The subscript t is introduced to show the dependence of the kernel on time t . This shows that the solution of the Cauchy problem (3) can be interpreted as kernel smoothing.

When the differential operator $\mathcal{L} = \Delta$, the Laplace–Beltrami operator, the Cauchy problem (3) becomes an isotropic diffusion equation. For this particular case, K_t^∞ is called the *heat kernel* with bandwidth t [7], [9]. For an arbitrary cortical manifold, the basis functions ψ_j can be computed and the exact shape of heat kernel can be determined numerically. Although it can be done by setting up a huge finite element method [39], this is not a trivial numerical computation. A simpler approach is to use the first order approximation of the heat kernel for small bandwidth and iteratively apply it up to the desired bandwidth [9].

WFS can be reformulated as a kernel regression problem [17]. At each fixed point p , we estimate unknown signal ν (4) with smooth function $h \in L^2(\mathcal{M})$ by minimizing the integral of the weighted squared distance between f and h

$$\min_{h \in L^2(\mathcal{M})} \int_{\mathcal{M}} K_t(p, q) |f(q) - h(p)|^2 d\mu(q). \quad (11)$$

The minimizer of (11) is given by the following theorem.

Theorem 2:

$$\frac{\mathcal{F}_t^k[f](p)}{\int_{\mathcal{M}} K_t^k(p, q) d\mu(q)} = \arg \min_{h \in L^2(\mathcal{M})} \int_{\mathcal{M}} K_t^k(p, q) \times |f(q) - h(p)|^2 d\mu(q).$$

Proof: Since the integral is quadratic in h , the minimum exists and obtained when

$$\begin{aligned} \frac{\partial}{\partial h} \int_{\mathcal{M}} K_t^k(p, q) |f(q) - h(p)|^2 d\mu(q) \\ = -2 \int_{\mathcal{M}} K_t^k(p, q) [f(q) - h(p)] d\mu(q) = 0. \end{aligned}$$

Solving the equation, we obtain the result.

Theorem 2 shows WFS is the solution of a weighted least squares minimization problem.

When \mathcal{L} is the Laplace–Beltrami operator with $k = \infty$, the heat kernel K_t^∞ is a probability distribution in \mathcal{M} , i.e.,

$$\int_{\mathcal{M}} K_t^\infty(p, q) d\mu(p) = 1.$$

For this special case, Theorem 2 simplifies to

$$\mathcal{F}_t^\infty[f](p) = \arg \min_{h \in L^2(\mathcal{M})} \int_{\mathcal{M}} K_t^\infty(p, q) |f(q) - h(p)|^2 d\mu(q).$$

In minimizing the weighted least squares in Theorem 2, it is possible to restrict the function space $L^2(\mathcal{M})$ to a finite subspace that is more useful in numerical implementation. Let

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

be the subspace spanned by basis ψ_0, \dots, ψ_l . Then we have the following theorem.

Theorem 3: If $\int_{\mathcal{M}} K_t^k(p, q) d\mu(q) = 1$ and $l \leq k$, then

$$\mathcal{F}_t^l[f](p) = \arg \min_{h \in \mathcal{H}_l} \int_{\mathcal{M}} K_t^k(p, q) |f(q) - h(p)|^2 d\mu(q).$$

Proof: Let $h(p) = \sum_{j=0}^l \beta_j \psi_j(p) \in \mathcal{H}_l$. The integral is written as

$$I(\beta_0, \dots, \beta_l) = \int_{\mathcal{M}} K_t^k(p, q) \left| f(q) - \sum_{j=0}^l \beta_j \psi_j(p) \right|^2 d\mu(q).$$

Since the functional I is quadratic in β_0, \dots, β_l , the minimum exists and it is obtained when $\partial I / \partial \beta_{j'} = 0$ for all j' . By differentiating I and rearranging terms, we obtain

$$\sum_{j=0}^k e^{-\lambda_j t} \psi_j(p) \psi_{j'}(p) \int_{\mathcal{M}} f(q) \psi_j(q) d\mu(q) = \sum_{j=0}^l \beta_j \psi_j(p) \psi_{j'}(p).$$

Now integrate the equations respect to measure $\mu(p)$ and obtain

$$e^{-\lambda_{j'} t} \langle f, \psi_{j'} \rangle = \beta_{j'}.$$

If K_t^k is a probability distribution, this theorem holds. For any other symmetric positive definite kernel, it can be made to be a probability distribution by renormalizing it. So Theorem 3 can be applicable in wide variety of kernels.

B. Isotropic Diffusion on Unit Sphere

Let us apply the WFS theory to a unit sphere denoted by S^2 . Since algebraic surfaces provide basis functions in a close form, it is not necessary to construct numerical basis [39]. The WFS in S^2 is given by the solution of isotropic diffusion. The spherical parametrization of S^2 is given by the polar angle θ and the azimuthal angel φ

$$p = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) \quad (12)$$

with $p = (\theta, \varphi) \in [0, \pi] \otimes [0, 2\pi]$. The spherical Laplacian Δ corresponding to the parametrization (12) is given by

$$\Delta = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2}.$$

There are $2l + 1$ eigenfunctions $Y_{lm} (-l \leq m \leq l)$, corresponding to the same eigenvalue $\lambda_l = l(l + 1)$ satisfying

$$\Delta Y_{lm} = \lambda_l Y_{lm}.$$

Y_{lm} is called the *spherical harmonic* of degree l and order m [12], [50]. It is given explicitly as

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

where $c_{lm} = \sqrt{((2l+1)/2\pi)((l-|m|)!/(l+|m|)!)}$ and P_l^m is the associated Legendre polynomials of order m .

Unlike many previous imaging literatures on spherical harmonics that used the complex-valued spherical harmonics [5], [23], [27], [44], only real-valued spherical harmonics are used throughout the paper for convenience in setting up a real-valued stochastic model.

For $f, h \in L^2(S^2)$, we define the inner product as

$$\langle f, h \rangle = \int_{\varphi=0}^{2\pi} \int_{\theta=0}^{\pi} f(p) h(p) d\mu(p),$$

where Lebesgue measure $d\mu(p) = \sin \theta d\theta d\varphi$. Then with respect to the inner product, the spherical harmonics satisfies the orthonormal condition

$$\int_{S^2} Y_{ij}(p) Y_{lm}(p) d\mu(p) = \delta_{il} \delta_{jm},$$

where δ_{il} is the Kroneker's delta. The kernel K_t^k is given by

$$K_t^k(p, q) = \sum_{l=0}^k \sum_{m=-l}^l e^{-l(l+1)t} Y_{lm}(p) Y_{lm}(q). \quad (13)$$

The associated WFS is given by

$$\mathcal{F}_t^k[f](p) = \sum_{l=0}^k \sum_{m=-l}^l e^{-l(l+1)t} f_{lm} Y_{lm}(p)$$

with Fourier coefficient $f_{lm} = \langle f, Y_{lm} \rangle$. We will call this form of WFS as the *weighted-SPHARM*. The special case $\mathcal{F}_0^k[f]$ is the traditional SPHARM representation used in representing the Cartesian coordinates of anatomical boundaries [23], [27], [44]. Consider subspace

$$\mathcal{H}_k = \left\{ \sum_{l=0}^k \sum_{m=-l}^l \beta_{lm} Y_{lm} : \beta_{lm} \in \mathbb{R} \right\} \subset L^2(S^2)$$

which is spanned by up to the k -th degree spherical harmonics. Then the SPHARM satisfy the least squares minimization problem different from Theorem 2 and Theorem 3.

Theorem 4:

$$\mathcal{F}_0^k[f](p) = \arg \min_{h \in \mathcal{H}_k} \|f - h\|^2. \quad (14)$$

III. NUMERICAL IMPLEMENTATION

In constructing the WFS representation, all we need is estimating Fourier coefficients $f_{lm} = \langle f, Y_{lm} \rangle$. There are three major techniques for computing the Fourier coefficients. The first method numerically integrates the Fourier coefficients over a high resolution triangle mesh [7]. Although this approach is the simplest to implement numerically and more accurate, the computation is extremely slow, due to the brute force nature of the technique. The second method is based on the fast Fourier transform (FFT) [5], [27]. The drawback of FFT is the need for a predefined regular grid system so if the mesh topology is different for each subject as in the case of FreeSurfer [18], a time consuming interpolation is needed. The third method is based on solving a system of linear equations [23], [43], [44] that minimize the least squares problem in Theorem. This is the most widely used numerical technique in the SPHARM literature. However, the direct application of the least squares estimation is not desirable when the size of the linear equation is extremely large.

Let

$$h = \sum_{l=0}^k \sum_{m=-l}^l \beta_{lm} Y_{lm}(p) \in \mathcal{H}_k.$$

Given n nodes p_1, \dots, p_n in S^2 mesh, the discretization of (14) is given by

$$\|f - h\|^2 \approx \sum_{i=1}^n \left[f(p_i) - \sum_{l=0}^k \sum_{m=-l}^l \beta_{lm} Y_{lm}(p_i) \right]^2. \quad (15)$$

The minimum of (15) is obtained when

$$f(p_i) = \sum_{l=0}^k \sum_{m=-l}^l \beta_{lm} Y_{lm}(p_i) \quad (16)$$

all $1 \leq i \leq n$. The (16) is referred as the *normal equation* in statistical literatures. The normal equation is usually solved via a matrix inversion. Let

$$\mathbf{f} = (f(p_1), \dots, f(p_n))'$$

and

$$\boldsymbol{\beta}_l = (\beta_{l,-l}, \dots, \beta_{l,l})'.$$

Also let

$$\mathbf{Y}_l = \begin{bmatrix} Y_{l,-l}(p_1) & \cdots & Y_{l,l}(p_1) \\ \vdots & \ddots & \vdots \\ Y_{l,-l}(p_n) & \cdots & Y_{l,l}(p_n) \end{bmatrix}$$

be a $n \times (2l+1)$ submatrix consisting of the l -th degree spherical harmonics evaluated at each node p_i . Then (16) can be rewritten in the following matrix form:

$$\mathbf{f} = \mathbf{Y}\boldsymbol{\beta} \quad (17)$$

with the *design matrix* $\mathbf{Y} = [\mathbf{Y}_0, \mathbf{Y}_1, \dots, \mathbf{Y}_k]$ and unknown parameter vector $\boldsymbol{\beta} = (\boldsymbol{\beta}'_0, \dots, \boldsymbol{\beta}'_k)'$. The linear system is solved via

$$\boldsymbol{\beta} = (\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{f}. \quad (18)$$

The problem with this widely used formulation is that the size of the matrix \mathbf{Y} is $n \times (k+1)^2$, which becomes fairly large and may not fit in typical computer memory. So it becomes impractical to perform matrix operation (18) directly. This is true for many cortical surface extraction tools such as FreeSurfer [18] that produces no less than $n > 100\,000$ nodes for each hemisphere. This computational bottleneck can be overcome by breaking the least squares problem in the subspace \mathcal{H}_k into smaller subspaces using the *iterative residual fitting* (IRF) algorithm [43]. The IRF algorithm was first introduced in [43] although the correctness of the algorithm was not given. In this paper, we present Theorem 5 that proves the correctness of the IRF for the first time. The IRF algorithm can be also used in estimating SPHARM coefficients by letting the bandwidth $t = 0$ in the algorithm.

A. Iterative Residual Fitting (IRF) Algorithm

Decompose the subspace \mathcal{H}_k into smaller subspaces as the direct sum

$$\mathcal{H}_k = \mathcal{I}_0 \oplus \mathcal{I}_1 \cdots \oplus \mathcal{I}_k$$

where subspace

$$\mathcal{I}_l = \left\{ \sum_{m=-l}^l \beta_{lm} Y_{lm}(p) : \beta_{lm} \in \mathbb{R} \right\}$$

is spanned by the l -th degree spherical harmonics only. Then the IRF algorithm estimates the Fourier coefficients $\boldsymbol{\beta}_l$ in each subspace \mathcal{I}_l iteratively from increasing the degree from 0 to k . Suppose we estimated the coefficients $\boldsymbol{\beta}_0, \dots, \boldsymbol{\beta}_{l-1}$ up to degree $l-1$ somehow. Then the residual vector \mathbf{r}_{l-1} based on this estimation is given by

$$\mathbf{r}_{l-1} = \mathbf{f} - \sum_{j=0}^{l-1} e^{-j(j+1)t} \mathbf{Y}_j \boldsymbol{\beta}_j. \quad (19)$$

The components of the residual vector \mathbf{r}_{l-1} are identical so we denote all of them as r_{l-1} . At the next degree l , we estimate the coefficients $\boldsymbol{\beta}_l$ by minimizing the difference between the residual r_{l-1} and $\sum_{m=-l}^l \beta_{lm} Y_{lm} \in \mathcal{I}_l$. This is formally stated as the following theorem.

Theorem 5:

$$\sum_{m=-l}^l f_{lm} Y_{lm} = \arg \min_{h \in \mathcal{I}_l} \|r_{l-1} - h\|^2. \quad (20)$$

Proof: Let $h = \sum_{m=-l}^l \beta_{lm} Y_{lm}(p) \in \mathcal{I}_l$. The squared norm is denoted as

$$I(\beta_{l,-l}, \dots, \beta_{l,l}) = \int_{S^2} \left[r_{l-1}(p) - \sum_{m=-l}^l \beta_{lm} Y_{lm}(p) \right]^2 d\mu(p).$$

Since I is quadratic in β_{lm} 's, the minimum is obtained when $\partial I / \partial \beta_{lm'} = 0$ for all m' . By differentiating I with respect to $\beta_{lm'}$ and letting it equal to zero, we obtain

$$\begin{aligned} \int_{S^2} Y_{lm'}(p) \left[r_{l-1}(p) - \sum_{m=-l}^l \beta_{lm} Y_{lm}(p) \right] d\mu(p) \\ = \int_{S^2} Y_{lm'}(p) r_{l-1}(p) d\mu(p) - \beta_{lm'}. \end{aligned} \quad (21)$$

From (19), we have

$$r_{l-1} = f - \sum_{j=0}^{l-1} \sum_{m=-j}^j e^{-j(j+1)t} \beta_{jm} Y_{jm}$$

which is a linear combination of spherical harmonics up to $(l-1)$ -th degree so it is orthonormal to Y_{lm} . Then the first term in (21) simplifies to

$$\int_{S^2} Y_{lm'}(p) r_{l-1}(p) d\mu(p) = \langle f, Y_{lm'} \rangle.$$

Theorem 5 proves that the correctness of the IRF algorithm. Then the discretization and the optimization is based on the normal equation approach (14)

$$\beta_l = (\mathbf{Y}_l \mathbf{Y}_l)^{-1} \mathbf{Y}_l' \mathbf{r}_{l-1}.$$

Summarizing the results, the IRF algorithm is given below.

Algorithm 1: Iterative Residual Fitting (IRF)

- Step 1) Let $l = 0$.
 Step 2) $\beta_0 \leftarrow (\mathbf{Y}_0 \mathbf{Y}_0)^{-1} \mathbf{Y}_0' \mathbf{f}$.
 Step 3) $l \leftarrow l + 1$.
 Step 4) $\mathbf{r}_{l-1} \leftarrow \mathbf{f} - \sum_{j=0}^{l-1} e^{-j(j+1)t} \mathbf{Y}_j \beta_j$.
 Step 5) $\beta_l \leftarrow (\mathbf{Y}_l \mathbf{Y}_l)^{-1} \mathbf{Y}_l' \mathbf{r}_{l-1}$.
 Step 6) If $l < k$, go to step 3.

B. Automatic Optimal Degree Selection in IRF

The IRF algorithm hierarchically builds the WFS from lower to higher degree. In most previous SPHARM literature [5], [23], [24], [27], [43], [44], the issue of the optimal degree has not been addressed. In SPHARM, the degree is simply selected based on

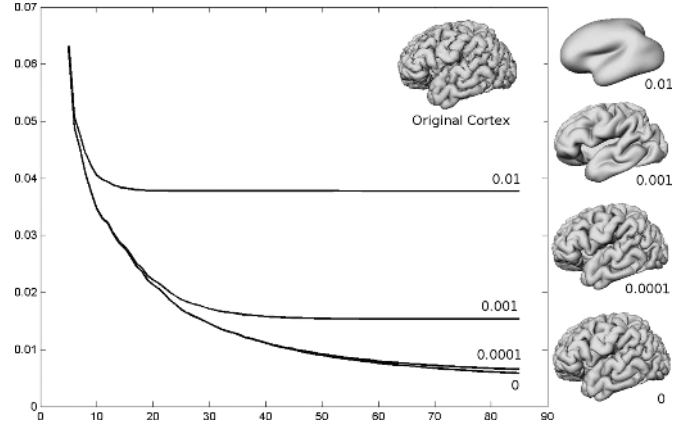


Fig. 1. Plots of RMSE over degree for bandwidths $t = 0.01, 0.001, 0.0001, 0$. Smoothed outer surfaces are at the 85th degree showing the bandwidth controls the amount of smoothing in representing the cortex. Bandwidth $t = 0$ corresponds to the traditional SPHARM. As $t \rightarrow 0$, the WFS converges to SPHARM.

a prespecified error bound that depends on the size of anatomical structure. We present a statistical framework for automatically determining the optimal degree that does not depend on the size of anatomical structure.

Although increasing the degree of WFS increases the goodness-of-fit, it also increases the number of coefficients to be estimated quadratically. So it is necessary to find the optimal degree where the goodness-of-fit and the number of parameters balance out.

The Fourier coefficients f_{lm} can be modeled to follow independent normal distribution $N(\mu_{lm}, \sigma_l^2)$. It is natural to assume an equal variance within the same degree. This assumption is equivalent to the following $(k-1)$ -th degree model

$$f(p_i) = \sum_{l=0}^{k-1} \sum_{m=-l}^l e^{-\lambda(\lambda+1)t} \mu_{lm} Y_{lm}(p_i) + \epsilon(p_i) \quad (22)$$

where ϵ is a zero mean isotropic Gaussian random field. Then at each iteration, we test if adding the k -th degree terms in the $(k-1)$ -th degree model is statistically significant by testing the null hypothesis

$$H_0 : \mu_{km} = 0 \text{ for } |m| \leq k.$$

The test statistic is constructed from the sum of squared errors (SSE). Let the k th degree *sum of squared errors* (SSE) be

$$\text{SSE}_k = \sum_{i=1}^n r_k^2(p_i).$$

The plot of the *root mean squared errors* (RMSE), $\sqrt{\text{SSE}_k/n}$ for varying degree $5 \leq k \leq 85$ is shown in Fig. 1. As the degree k increases, the root mean squared errors keep decreasing until it flattens out. So it is reasonable to stop the iteration when the decrease in error is no longer significant. Under H_0 , the test statistic is

$$F = \frac{(\text{SSE}_{k-1} - \text{SSE}_k)/(2k+1)}{\text{SSE}_{k-1}/(n - (k+1)^2)} \sim F_{2k+1, n-(k+1)^2}$$

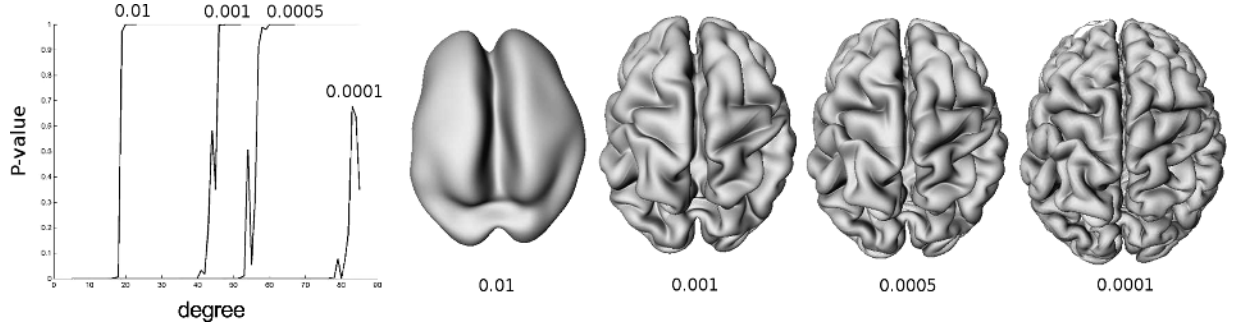


Fig. 2. Automatic degree selection in the IRF algorithm. For each bandwidth t , the optimal degree is automatically selected by checking if adding an additional degree will be statistical significant. Outer cortical surfaces are the results of the optimal selection procedure. The optimal degrees are $k = 18(t = 0.01)$, $k = 42(t = 0.001)$, $k = 52(t = 0.0005)$, and $k = 78(t = 0.0001)$. For our study, $t = 0.0001$ and the corresponding degree $k = 78$ is used through the paper.

the F -distribution with $2k + 1$ and $n - (k + 1)^2$ degrees-of-freedom. We compute the F statistic at each degree and stop the IRF procedure if the corresponding P -value first becomes bigger than the prespecified significance α (traditionally at 0.01 or 0.05 and $\alpha = 0.01$ in this study). For bandwidth $t = 0.0001$, the optimal degree is determined to be $k = 78$ (Fig. 2).

C. Computational Cost and Numerical Accuracy

The computation is done with MATLAB 6.1 in a Pentium 1.4-GHz Intel M -processor laptop with 512-MB memory. The complete MATLAB codes are freely available.¹ The IRF algorithm requires computing and storing the SPHARM bases in a hard drive. Sampling the SPHARM bases up to 85 degree at approximately 40 000 points on a unit sphere requires 2.4 GB of space and 16 min of computation. The SPHARM bases at each degree are stored as a single file and loaded into the computer at each degree in the IRF algorithm. For up to $k = 78$ degree, there are total $3(k+1)^2 = 18,723$ unknown Fourier coefficients corresponding to the three Cartesian coordinates of a cortical surface. The average computational times for estimating all the Fourier coefficients are 354, 345, and 313 s, respectively, for 1, 24, and 48 cortical surfaces. The average computational time decreases as the number of cortical surfaces increases since the SPHARM bases are recycled for all surfaces at a given degree.

The numerical implementation of WFS is validated against the known analytical value. From the orthonormality of SPHARM, for $l \leq k$

$$e^{l(l+1)t} \mathcal{F}_t^k[Y_{lm}](p) = Y_{lm}(p).$$

WFS of the spherical harmonic Y_{lm} multiplied by $e^{l(l+1)t}$ should theoretically results in the original spherical harmonics. This exact analytical relation gives a criteria for numerical accuracy. The difference between the analytical value and the WFS representation gives the measure of numerical accuracy. Table I shows the result for selective l and m with various bandwidth t . The fourth column shows the mean absolute error computed over all mesh vertices as

$$\frac{1}{n} \sum_{i=1}^n |Y_{lm}(p_i) - e^{l(l+1)t} \mathcal{F}_t^k[Y_{lm}](p_i)|.$$

¹<http://www.stat.wisc.edu/~mchung/softwares/weighted-SPHARM/weighted-SPHARM.html>

TABLE I
ACCURACY OF THE WFS REPRESENTATION

degree l	order m	bandwidth t	mean error	f_{lm}
18	17	0	0.0077	0.9979
18	17	0.0001	0.0078	0.9979
18	17	0.0005	0.0083	0.9981
18	17	0.01	0.0575	0.9995
42	41	0	0.0064	0.9977
42	41	0.001	0.0126	0.9992
52	51	0	0.0066	0.9972
52	51	0.0005	0.0101	0.9988
78	77	0	0.0060	0.9973
78	77	0.0001	0.0068	0.9984

As expected, the mean absolute error decreases as the degree increases. For the 78th degree with $t = 0.0001$, the error is smaller than two decimal places. We have also estimated the Fourier coefficient $f_{lm} = \langle Y_{lm}, Y_{lm} \rangle$ to check if it is 1 in the fifth column. The estimation is accurate up to two decimal places for all degrees. A similar validation technique was introduced in [7] for determining the accuracy of a numerical integration.

We have also checked the ringing artifacts (Gibbs phenomenon) [22] in our numerical implementation. In both the traditional SPHARM and WFS, continuous measurements will have rapidly decaying Fourier coefficients so they converge faster. For continuous measurements such as the Cartesian coordinates and cortical thickness, we do not expect significant ringing artifacts for fairly large degree of harmonics. Note that we are using sufficiently high degree up to 78th harmonics, which corresponds to 6241 independent basis functions.

On the other hand, discontinuous measurements will have slowly decaying Fourier coefficients and thus the representations converge slowly. The ringing artifacts would be expected for the SPHARM representation as demonstrated in Fig. 5. In the figure, discontinuous measurements are constructed as a step function of value 1 in the circular band $(1/8) < \theta < (1/4)$ and 0 outside of the band. The SPHARM representation of the step function resulted in significant ringing artifacts even for fairly high degrees up to 78. In comparison, the WFS representation of the step function does not exhibit any serious ringing artifacts as seen in Fig. 5. The superior performance of WFS can be easily explained in terms of convergence. The WFS representation additionally weights Fourier coefficients with exponentially decaying weights, which contributes more rapid convergence even for discontinuous measurements. This robustness of WFS is also related to the fact that it is a PDE-based data

smoothing technique while the traditional SPHARM is more of interpolation or reconstruction technique.

D. Computing FWHM

Since WFS is a kernel smoothing method, it is useful to know the full-width at the half-maximum (FWHM) of the underlying kernel. The computed FWHM is later used in the random field based multiple comparison corrections [9], [10], [52], [53]. Computing the FWHM of the heat kernel used in WFS is not trivial since there is no known close form expression for FWHM as a function of bandwidth t . Therefore, FWHM is computed numerically.

For $p, q, r \in S^2$, let us define the Cartesian inner product \cdot as $p \cdot q = \cos(\theta)$, where θ is an angle between p and q . The heat kernel (13) is symmetric along the geodesic circle. If $p \cdot q = p \cdot r$, we have $K_t^k(p, q) = K_t^k(p, r)$. This property can be used to simplify the expansion (13) using the harmonic addition theorem [26], [50].

Theorem 6: (Harmonic Addition Theorem):

$$\sum_{m=-l}^l Y_{lm}(p)Y_{lm}(q) = \frac{2l+1}{4\pi} P_l^0(p \cdot q). \quad (23)$$

Proof: Fix the azimuthal angel $\varphi = 0$ and p be the north pole, i.e. $p = (0, 0, 1)$. Now by varying $q = (\sin \theta, 0, \cos \theta)$ for $0 \leq \theta \leq \pi$, we have $Y_{lm}(\theta, \varphi) = 0$ if $m \neq 0$. Then we have

$$\sum_{m=-l}^l Y_{lm}(p)Y_{lm}(q) = Y_{l0}(p)Y_{l0}(q) = \frac{2l+1}{4\pi} P_l^0(1)P_l^0(\cos \theta).$$

Note that $P_l^0(p) = 1$. This implies that the sum of product of the spherical harmonics is a function of the inner product between p and q only. From symmetry, if we rotate p back to the original position from the north pole, the same result should hold.

Using the harmonic addition theorem, we simplify the heat kernel in the following theorem.

Theorem 7: For any $p, q \in S^2$

$$K_t^k(p, q) = \sum_{l=0}^k \frac{2l+1}{4\pi} e^{-l(l+1)t} P_l^0(p \cdot q). \quad (24)$$

Theorem 7 is used to plot the shape of the heat kernel by fixing p to be the north pole and by varying $\theta = \cos^{-1}(p \cdot q)$ (Fig. 3). Similar result is also given in [5]. The maximum of the kernel is obtained at $\theta = 0$. Then the FWHM is solved numerically for θ in

$$\frac{1}{2} \sum_{l=0}^k \frac{2l+1}{4\pi} e^{-l(l+1)t} = \sum_{l=0}^k \frac{2l+1}{4\pi} e^{-l(l+1)t} P_l^0(\cos \theta).$$

The FWHM is then 2θ . Fig. 4 shows the nonlinear relationship between bandwidth t and the corresponding FWHM. When $t = 0.0001$, the corresponding FWHM is 0.0597. This is the FWHM we have used in the subsequent cortical thickness analysis.

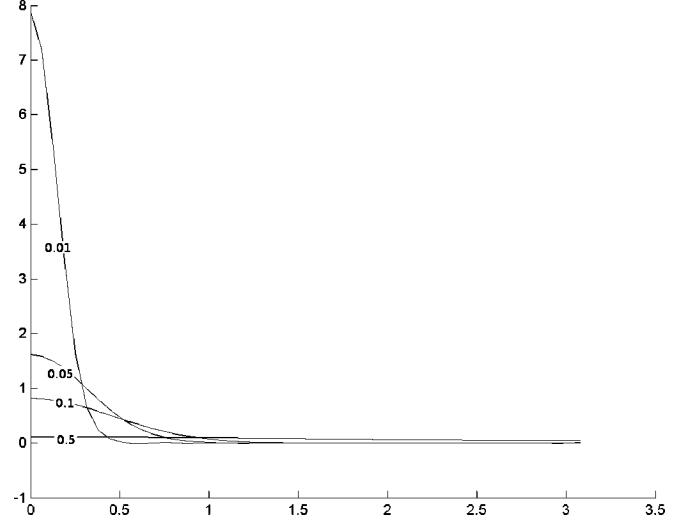


Fig. 3. Shape of heat kernel with different bandwidth $t = 0.01, 0.05, 0.1, 0.5$. Horizontal axis is the θ value from the north pole ($\theta = 0$) to the south pole ($\theta = \pi$). Weighting scheme used in WFS follows the shape of the heat kernel weighting closer measurements heavily.

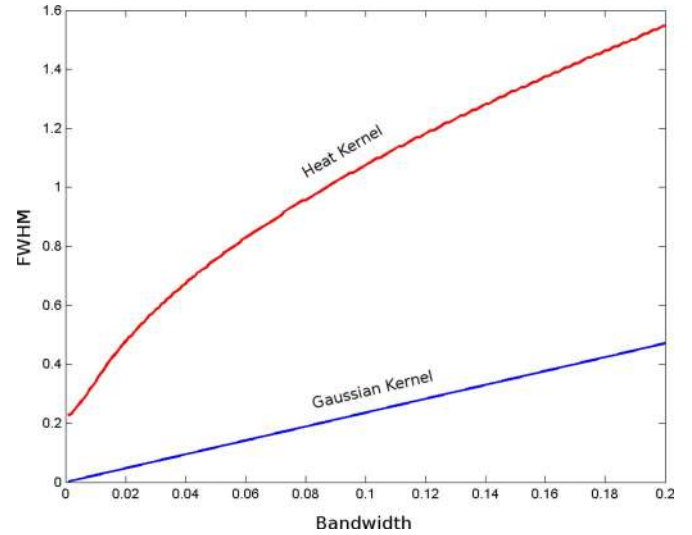


Fig. 4. Plot of FWHM (vertical) over bandwidth t (horizontal) for both heat kernel and Gaussian kernel. FWHM has to be numerically estimated in the case of the heat kernel. Numerically computed FWHM is used in the random field theory based multiple comparison correction.

IV. APPLICATION TO CORTICAL MORPHOMETRY

In this section, we present applications of WFS in localizing the regions of abnormal amount of gray matter in autistic subjects. Surface normalization issue is discussed in detail along with a stochastic model for WFS.

A. Data Set

The $n_1 = 12$ high functioning autistic (HFA) and $n_2 = 12$ normal control (NC) male subjects were screened to be right-handed. The autistic subjects were diagnosed via the Autism Diagnostic Interview—Revised (ADI-R) by a trained and certified

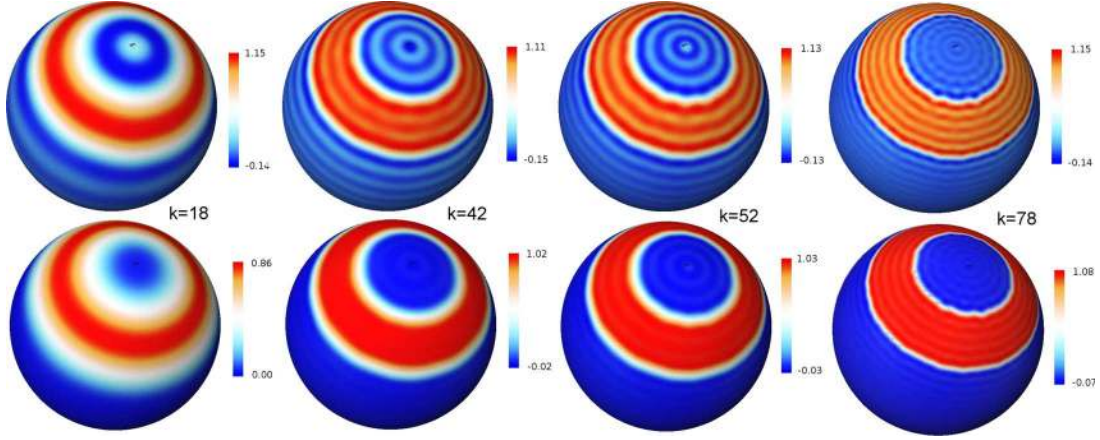


Fig. 5. Ringing artifacts (Gibbs phenomenon) for SPHARM and WFS. Traditional SPHARM and WFS representations are performed on the discontinuous measurements, which are defined as 1 in region $(1/8) < \theta < (1/4)$ and 0 in other regions. First row shows severe ringing artifacts in the traditional SPHARM representation. Second row shows the negligible ringing effect in the WFS representation. Columns are $k = 18, 42, 52, 78$ degree representations.

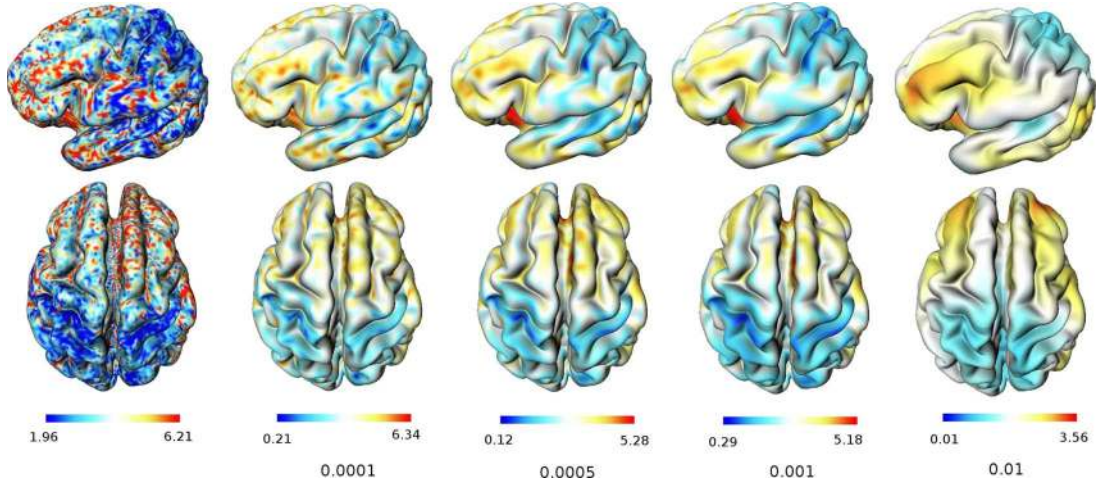


Fig. 6. Cortical thickness measures projected onto an average surface for better visualization. As the bandwidth increases from $t = 0.0001$ to $t = 0.01$, the amount of smoothing also increases. First image shows the cortical thickness obtained from the traditional deformable surface algorithm [9], [10], [32].

psychologist at the Waisman Center at the University of Wisconsin-Madison [14]. Age distributions for HFA and NC were compatible at 15.93 ± 4.71 and 17.08 ± 2.78 , respectively.

High resolution anatomical magnetic resonance images (MRI) were obtained using a 3-T GE SIGNA scanner with a quadrature head RF coil. A 3-D, spoiled gradient-echo (SPGR) pulse sequence was used to generate T_1 -weighted images. Image intensity nonuniformity was corrected using the nonparametric nonuniform intensity normalization method [46] and then the image was spatially normalized into the Montreal neurological institute (MNI) stereotaxic space using a global affine transformation [11]. Afterwards, an automatic tissue-segmentation algorithm based on a supervised artificial neural network classifier was used to classify each voxel as cerebrospinal fluid (CSF), gray matter, or white matter [31]. Subsequently, a deformable surface algorithm [32] was used to generate the outer and the inner cortical meshes.

B. Stochastic Model

Let \mathcal{M}_o and \mathcal{M}_i be the outer (pial) and inner surfaces of the brain, respectively. The unit sphere S^2 is realized as a triangle

mesh and deformed to match the outer and inner surfaces in such a way that anatomical homology and the topological connectivity of meshes are preserved [32]. The cortical surfaces can be assumed to be smooth 2-D Riemannian manifolds parameterized by two parameters [15], [29]. Based on the deformable algorithm [32] that establishes the homology between the S^2 mesh and the outer cortical surface, the Cartesian coordinates of the mapping are discretely parameterized by the spherical parametrization (12) as

$$v = (v_1(p), v_2(p), v_3(p)).$$

The inner surface is parameterized similarly as

$$w = (w_1(p), w_2(p), w_3(p)).$$

These discrete coordinate functions are further smoothed by the WFS

$$v_i(p) = \sum_{l=0}^k \sum_{m=-l}^l e^{-l(l+1)t} f_{lm}^i Y_{lm}(p). \quad (25)$$

We model v_i stochastically as (22) by assuming f_{lm}^i to follow independent normal distribution $N(\mu_{lm}^i, \sigma_l^2)$ for coordinate i , degree l , and order m . This assumption is equivalent to modeling v_i as the sum of signal plus noise

$$v_i(p) = \sum_{l=0}^k \sum_{m=-l}^l e^{-l(l+1)t} \mu_{lm}^i Y_{lm}(p) + \epsilon_i(p)$$

where ϵ_i is a zero mean Gaussian random field with a certain isotropic covariance function. A similar stochastic modeling approach has been used in [34] where the canonical expansion of a Gaussian random field is used to model the component of a deformation field.

The mean and the variance functions of the surface are given by

$$\begin{aligned} \mathbb{E}v_i(p) &= \sum_{l=0}^k \sum_{m=-l}^l e^{-l(l+1)t} \mu_{lm}^i Y_{lm}(p), \\ \mathbb{V}v_i(p) &= \sum_{l=0}^k \sum_{m=-l}^l e^{-2l(l+1)t} \sigma_l^2 Y_{lm}^2(p). \end{aligned} \quad (26)$$

The total variability of the surface is then measured by

$$\int_{S^2} \mathbb{V}v_i d\mu(p) = \sum_{l=0}^k \sum_{m=-l}^l e^{-2l(l+1)t} \sigma_l^2$$

indicating the increase of smoothing bandwidth decreases the total variability. If

$$v_{ij}(\theta, \varphi) = \sum_{l=0}^k \sum_{m=-l}^l e^{-l(l+1)t} f_{lm}^{ij} Y_{lm}(\theta, \varphi) \quad (27)$$

is the WFS for the j th subject ($1 \leq j \leq s$), the unknown parameters μ_{lm}^i and σ_l^2 are estimated as the sample mean and the sample variance

$$\begin{aligned} \widehat{\mu_{lm}^i} &= \frac{1}{s} \sum_{j=1}^s f_{lm}^{ij}, \\ \widehat{\sigma_l^2} &= \frac{1}{(2l+1)(s-1)} \sum_{m=-l}^l \sum_{j=1}^s (f_{lm}^{ij} - \widehat{\mu_{lm}^i})^2. \end{aligned} \quad (28)$$

The inner surface is stochastically modeled similarly as

$$w_i(p) = \sum_{l=0}^k \sum_{m=-l}^l e^{-l(l+1)t} g_{lm}^i Y_{lm}(p). \quad (29)$$

C. Surface Normalization

Previously, cortical surface normalization was performed by minimizing an objective function that measures the global fit of two surfaces while maximizing the smoothness of the deformation in such a way that the gyral patterns are matched smoothly [9], [42], [49]. In the WFS representation, the surface normalization is straightforward and does not require any sort of optimizations explicitly.

The MRIs were spatially normalized into the Montreal Neurological Institute (MNI) stereotaxic space using a global affine transform [11], which gives the 3-D alignment of brain. Then an additional surface alignment was done during the cortical surface extraction process via the anatomic segmentation using the proximities (ASP) algorithm [32]. The algorithm generates 40 962 vertices and 81 920 triangles with the identical mesh topology for all subjects. The vertices indexed identically on two cortical meshes will have a very close anatomic homology and this defines the surface alignment [10], [32]. This provides the same spherical parameterization at identically indexed vertices across different cortical surfaces. Our approach avoids the surface alignment done by coinciding the first order ellipsoid meridian and equator in the SPHARM-correspondence approach [23], [47]. Surface meshes obtained from other segmentation techniques such as FreeSurfer [18] may require the SPHARM-correspondence approach. After this initial surface alignment, the more refined WFS-correspondence is established.

Given two WFS surfaces v_{i1} and v_{i2} , the displacement field d_i that minimizes the integral of the squared errors between deformed surface $d_i(v_{i1})$ and v_{i2} is simply given by the following theorem.

Theorem 8:

$$v_{i2} - v_{i1} = \arg \min_{d_i \in \mathcal{H}_k} \int_{S^2} [v_{i1} + d_i(v_{i1}) - v_{i2}]^2 d\mu(p).$$

Proof: If $d_i \in \mathcal{H}_k$, the deformation is of the form

$$v_{i1} + d_i(v_{i1}) = \sum_{l=0}^k \sum_{m=-l}^l [e^{-l(l+1)t} f_{lm}^{i1} + \beta_{lm}^i] Y_{lm}$$

for unknown β_{lm}^i , which need to be determined. Substituting the term in the integral, we obtain the expression

$$I = \int_{S^2} \left[\sum_{l=0}^k \sum_{m=-l}^l \alpha_{lm}^i Y_{lm} \right]^2 d\mu(p)$$

where $\alpha_{lm}^i = e^{-l(l+1)t} f_{lm}^{i1} + \beta_{lm}^i - e^{-l(l+1)t} f_{lm}^{i2}$. From the orthonormality of spherical harmonics, the expression is simplified as

$$I = \sum_{l=0}^k \sum_{m=-l}^l \alpha_{lm}^i{}^2.$$

This is minimum if all α_{lm}^i vanishes and we obtain the result.

Theorem 8 shows that the optimal displacement in the least squares sense is obtained by simply taking the difference between two WFS representations. Unlike other surface registration methods used in warping surfaces between subjects [9], [42], [49], it is not necessary to consider an additional cost function that guarantees the smoothness of the displacement field since the displacement field $v_{i2} - v_{i1}$ is already a linear combination of smooth basis functions. Based on this idea, we normalize WFS surfaces. We will refer to the surface correspondence obtained by Theorem 8 as the *WFS-correspondence*.

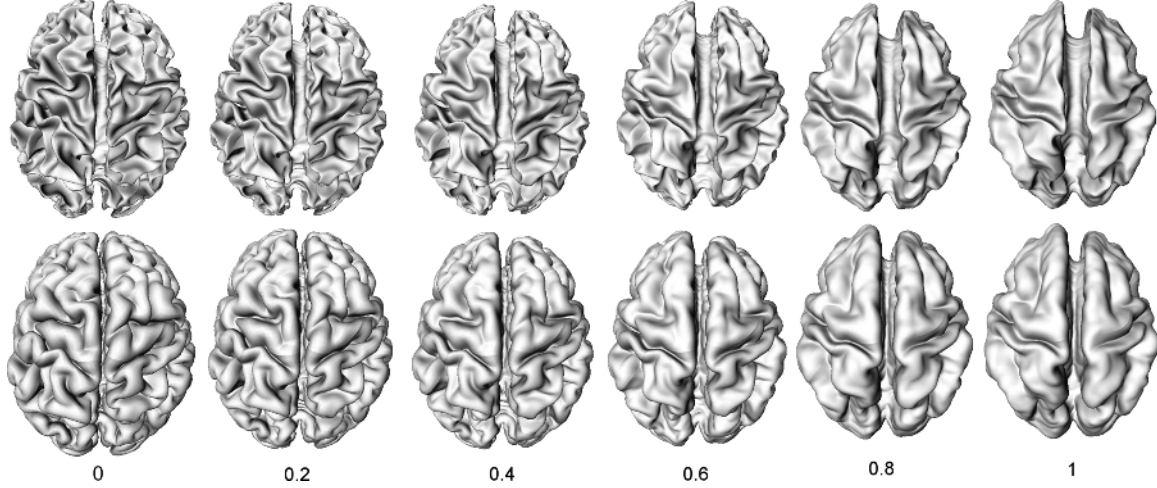


Fig. 7. Multiscale representation of surface registration toward the average template. Top is the inner surface while the bottom is the outer surface. $\alpha = 0$ is the surface of one particular subject while $\alpha = 1$ is the average surface of 24 subject. Scale space $\alpha \in [0, 1]$ is searched for the maximal discrepancy between two groups for increased statistical sensitivity.

Let \bar{v}_i be the mean surface obtained by replacing μ_{lm}^i in (26) with the sample mean (28). Fig. 7 shows the mean surface for 24 subjects used in the study. The mean surface serves as a template for statistical analysis later. For subject j , the displacement from surface v_{ij} to the template is

$$\Delta v_{ij} = \bar{v}_i - v_{ij}.$$

Consider the trajectory of the deformation from v_{ij} to the template \bar{v}_i parameterized by $\alpha \in [0, 1]$

$$\mathbf{v}_{ij}(\alpha) = v_{ij} + \alpha \Delta v_{ij} = (1 - \alpha)v_{ij} + \alpha \bar{v}_i. \quad (30)$$

When $\alpha = 0$, $\mathbf{v}_{ij}(\alpha)$ is the j th subject surface, while when $\alpha = 1$, it is the template. The parameter α controls the amount of registration from the coarse to fine scale toward the template. Fig. 7 shows $\mathbf{v}_i(\alpha)$ at 11 different scales between 0 and 1 with 0.1 increment for a single subject. The larger the value of α , the smaller the image registration variability across the subjects with respect to the template. This is shown from the total variability computed at each scale α :

$$\int_{S^2} \mathbb{V}(\mathbf{v}_{ij}) d\mu(p) = c(\alpha) \sum_{l=0}^k \sum_{m=-l}^l e^{-2l(l+1)t} \sigma_l^2$$

where

$$c(\alpha) = \frac{n-1}{n^2} \alpha^2 + \left(1 - \frac{n-1}{n} \alpha\right)^2$$

is decreasing for $0 \leq \alpha \leq 1$ (Fig. 8).

D. Gray Matter Density

Gray matter density is a 3-D measure defined as the probability of a particular voxel belonging to gray matter. It has been used in various anatomical studies: normal development [25],

[37], autism [8], depression [38], epilepsy [33], and Alzheimer's disease [28], [48]. There are many different techniques for obtaining the density depending on how it is defined. In VBM, it is modeled as a Gaussian mixture on tissue intensity values [2], [25]. In modulated-VBM [25], the density obtained from the standard VBM is rescaled by the Jacobian determinant of image registration to preserve the total amount of gray matter. This is related to the RAVENS (regional analysis of volumes examined in normalized space) approach [16]. In this study, we have avoided modulating the density with the Jacobian determinant. Since only the surface deformation is known, we do not have any information about the Jacobian determinant inside the gray matter regions.

Paus *et al.* modeled the density as a Bernoulli random variable taking value 1 inside the gray matter segmentation and 0 outside the segmented regions [37]. In a slightly different formulation, Thompson *et al.* computed the density as the fraction of gray matter within a ball of radius 15 mm along a cortical surface [48]. This approach is equivalent to convoluting the binary mask of the gray matter with a uniform probability distribution of radius 15 mm and interpolating voxel values to the cortical surface mesh. This equivalence relation is the basis of how we project a 3-D density map to a 2-D cortical surface and compare them with cortical thickness in our study.

In this study, the gray matter density was constructed using the 3-D Euclidian distance map of the surfaces. For the outer surface \mathcal{M}_o , the distance map at each voxel x is defined as

$$\text{dist}_o(x) = \min_{y \in \mathcal{M}_o} \|x - y\|$$

where $\|\cdot\|$ is the Euclidian norm. The minimum is found using the nearest neighbor search algorithm on an optimized k-D tree [20]. Similarly we denote the distance map for the inner surface \mathcal{M}_i as $\text{dist}_i(x)$. Then the average distance map is defined as

$$\text{dist}(x) = \frac{\text{dist}_o(x) + \text{dist}_i(x)}{2}.$$

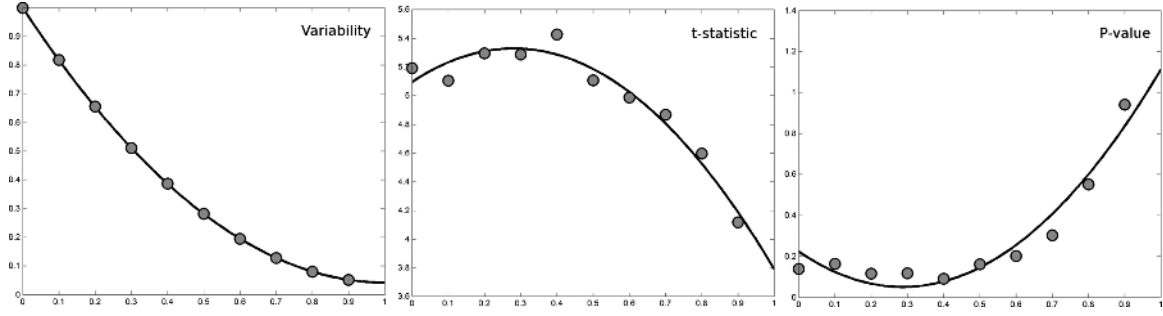


Fig. 8. Left: Plot of image registration variability $c(\alpha)$. Middle: Plot of the maximum of T statistic at each scale. Right: Plot of corrected P -value corresponding to the maximum of T statistic. At $\alpha = 0.4$, the minimum P -value of less than 0.1 is obtained so we choose $\alpha = 0.4$ to be the optimal scale that separates the two groups. Lines are the best fitting quadratic curve in the least squares sense.

The average distance map for a subject is shown in Fig. 9. The minimum of the average distance is always obtained in the middle of the outer and the inner surfaces, where the probability of a voxel belong to the gray matter class should be the highest. Then we define the gray matter density as

$$\text{density}(x) = \exp \left[-\frac{\text{dist}^2(x)}{2\rho^2} \right] \quad (31)$$

where parameter ρ^2 controls the spread of density ($\rho^2 = 3$ in our study). The gray matter density is always between 0 and 1 and it obtains its maximum in the interior of the gray matter region, where the average distance map obtains the minimum. The density map is further convoluted with the 3-D Gaussian kernel K with 10 mm FWHM to increase the smoothness and normality of data [2], [8] (Fig. 9). The smoothed density map $K * \text{density}(x)$ is stochastically modeled as a Gaussian random field.

Afterwards, the two sample t -test statistic T with the equal variance assumption is computed on the convoluted gray matter density maps at each voxel [8]. The resulting T random field $T(x)$ is distributed as a student t distribution with $\nu = n_1 + n_2 - 2$ degrees-of-freedom at each voxel x . Based on the random field theory [51]–[53], the test statistic, which accounts for the multiple comparison correction, is the maxima of T field over the gray matter \mathcal{M}_g . The corresponding *corrected P-value* is computed using the following formula:

$$P \left(\sup_{x \in \mathcal{M}_g} T(x) > h \right) \approx \frac{\text{Vol}(\mathcal{M}_g)}{\text{FWHM}^3} \frac{(4 \ln 2)^{3/2}}{(2\pi)^2} \times \left(\frac{\nu-1}{\nu} h^2 - 1 \right) \left(1 + \frac{h^2}{\nu} \right)^{-\frac{\nu-1}{2}}$$

where $\text{Vol}(\mathcal{M}_g) = 2.13 \times 10^5 \text{ mm}^3$ is the volume of the gray matter of the template. The gray matter volume is estimated by computing the volume bounded by the outer and inner meshes [10]. Restricting the search region from the whole brain volume to the gray matter boosts the signal detection power.

We performed the above procedure for each image registration scale from 0 to 1 at 0.1 increment. Increasing surface registration toward the template reduces the registration variability while increasing the gray density variability due to the misalignment of gray matter. So the maxima of T field will not be obtained at the two extremes of registration scale ($\alpha = 0, 1$) as

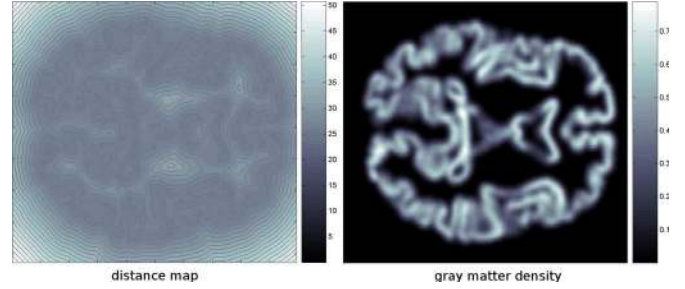


Fig. 9. Left: Contour plot of the average distance map in mm at scale $\alpha = 0$. Right: Gaussian kernel smoothing of the gray matter density map with 10 mm FWHM.

demonstrated in Fig. 8. To increase statistical sensitivity and specificity, it is necessary to find an optimal registration scale that provides the maximal discrepancy between the groups as measured by the maxima of T field. Hence, the minimum corrected P -value can be chosen as a criteria for determining the optimal scale. Our approach is similar to the scale-space search method [45], where the maxima of T field is searched over smoothing scale.

The maximum T statistic value and its corresponding corrected P -value at each scale are plotted in Fig. 8 showing that the optimal scale is obtained when $\alpha = 0.4$. At this scale, the maximum T -stat. was 5.43 while the minimum T -stat. was -5.04 . The random field theory based thresholding of $h = \pm 5.35$ gives the corrected P -value of 0.1. Fig. 10 shows the optimally constructed T -stat. map thresholded at ± 4.0 and interpolated into the nearest point in the cortical surface showing increased gray matter density in the localized areas of the autistic subjects.

E. Cortical Thickness

The previously available approaches for computing the cortical thickness in discrete triangle meshes produce noisy thickness measures [9], [18], [32]. So it is necessary to smooth the thickness measurements along the cortex via PDE based smoothing techniques [1], [6], [10]. On the other hand, the WFS provides smooth functional representation of the outer and inner surfaces so that the distance measures between the surfaces should be already smooth. Hence, the WFS avoids this additional step of thickness smoothing done in most of thickness analysis literature [9], [10]. It is not necessary to perform data smoothing in the WFS formulation.

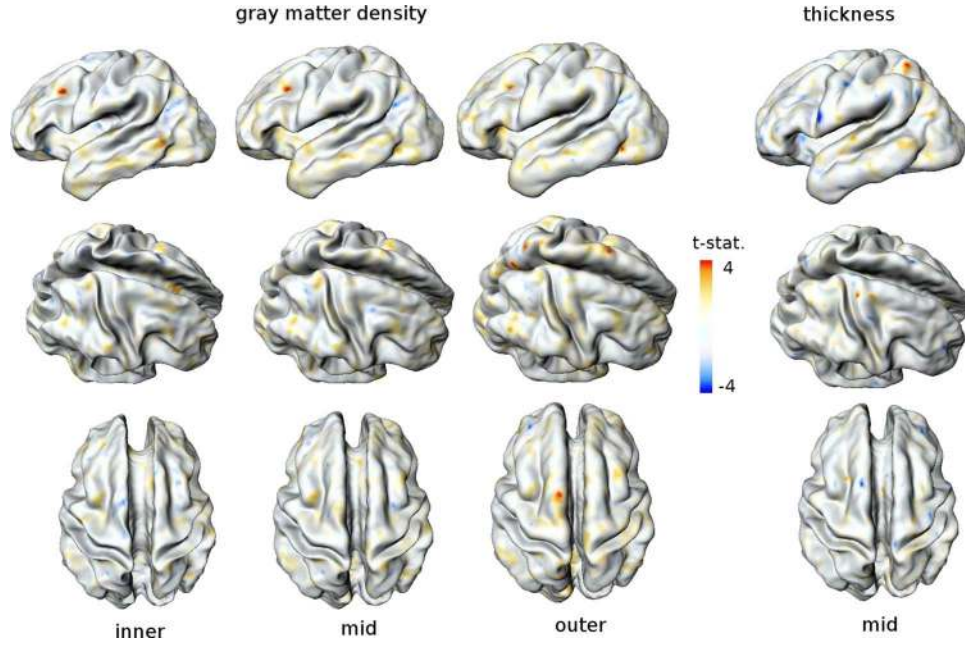


Fig. 10. T statistic map for gray matter density and cortical thickness thresholded at between -4 and 4 for the comparison purpose. In the gray matter density results, signals are mainly detected in either the inner surface or the outer surface but not in the middle surface. This indicates that density-based morphometries will mainly detect signal near tissue boundaries. Gray matter density and the thickness results do not overlap reconfirming that the thickness and gray matter density are not positively correlated.

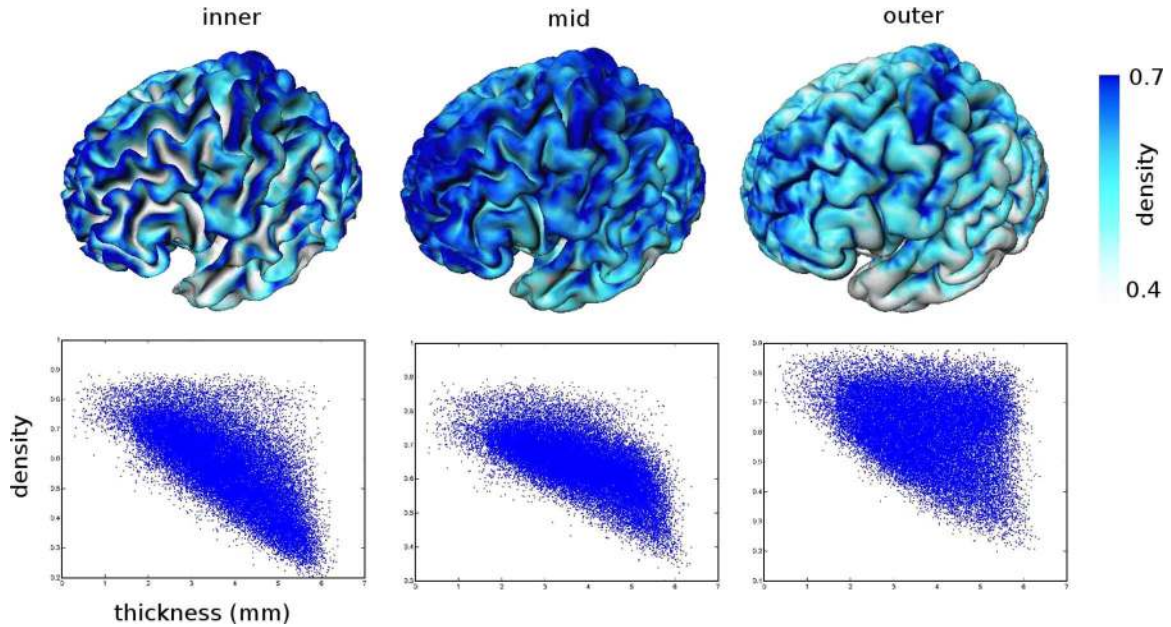


Fig. 11. Top: Gray matter density projected onto inner, middle, and outer surfaces. On the inner surface, the deep sulcal regions show the low density while the gyral ridges show high density. On the outer surface, this is opposite. The deep sulcal regions show high density while the gyral ridges show lower density. The middle surface shows high density. Bottom: Scatter plot of gray matter density over thickness. They show negative correlations.

Using Theorem 8 and the resulting WFS-correspondence, we establish the homology between the outer and the inner surfaces in the least squares fashion. For the outer surface (25) and the inner surface (29), the cortical thickness is defined to be the Euclidean distance between the WFS-correspondence. A similar approach has been proposed for measuring the closeness between two surfaces [23]; however, this is the first study using harmonics in defining the cortical thickness. Fig. 6 shows the comparison of cortical thickness computed from the tradi-

tional deformable surface algorithm [32] and the WFS-correspondence. The cortical thickness obtained from the traditional approach introduces a lot of triangle mesh noise into its estimation while the WFS-correspondence approach does not. The spatial smoothness of the thickness is controlled by the bandwidth t .

For the group comparison between the autistic and the normal control groups, a two sample t -test was performed. The corresponding corrected P -value was computed using the following

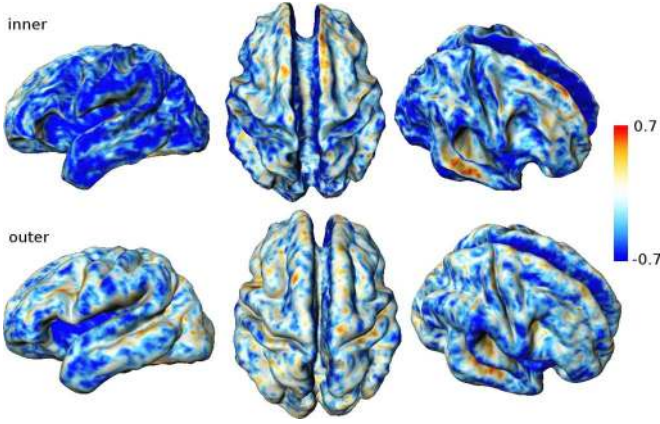


Fig. 12. Correlation of thickness and gray matter density for 24 subjects mapped on both the inner and the outer surfaces. Most of both inner and outer surfaces show negative correlation. Thicker cortical regions are less convoluted so the gray matter density tend to be lower.

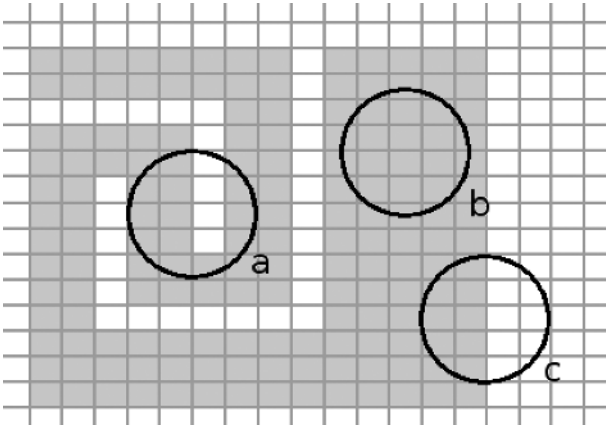


Fig. 13. Simple 2-D schematic showing the negative correlation between thickness and gray matter density. Gray colored pixels are gray matter. Black circle is the contour of heat kernel. There are more gray matter pixels in region (a) than region (c) although the thickness in region (c) is thicker than that of region (a). The gray matter density in the middle of the gray matter (b) is 1 for almost all subject indicating very small between-subject and between-group variability. Because of the small between-group variability, VBM does not usually detect signal in the middle of the gray matter. Most of significant signal detected in VBM is near the tissue boundary where the between-group variability is high.

formula:

$$P\left(\sup_{p \in S^2} T(p) > h\right) \approx \frac{1}{\text{FWHM}^2} \frac{(4 \ln 2)}{(2\pi)^{1/2}} \times \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\left(\frac{\nu}{2}\right)^{1/2} \Gamma\left(\frac{\nu}{2}\right)} h \left(1 + \frac{h^2}{\nu}\right)^{-\frac{\nu-1}{2}}$$

where FWHM is from the heat kernel used in the WFS. We used $\text{FWHM} = 0.0597$ corresponding to bandwidth $t = 0.0001$. However, we did not detect any statistically significant region at $\alpha = 0.05$ level. The minimum T -stat. was -4.73 while the maximum T -stat. was 4.83 . Fig. 10 shows the T -stat. map thresholded at ± 4 .

F. Comparing Cortical Thickness and Gray Matter Density

Most morphometric studies [2], [3], [9], [10], [18], [25] have never compared cortical thickness and gray matter density to-

gether so it is not clear if the two anatomical indices are positively correlated. Comparing the SPM of density and thickness in Fig. 10, no statistically significant regions overlap. Since both metrics have been assumed to be the indicators of the amount of gray matter, the result is paradoxical. We have correlated these two metrics within a subject (Fig. 11) and across subjects (Fig. 12). Surprisingly, the scatter plot in Fig. 11 shows negative correlation. By assigning the density value of a voxel that contains a vertex of a cortical mesh to the vertex, we can project the gray matter density onto inner, middle and outer surfaces. The middle surface is obtained by averaging the inner and the outer surfaces in the WFS-correspondence. Fig. 12 shows negatively correlated scatter plots. Fig. 12 shows the complex pattern of nonuniformity of density. On the outer surface, deep sulci have higher density compared to gyri while on the inner surface, the pattern is opposite. The middle surface shows higher density compared to the outer and the inner surfaces as expected. These complex patterns of the nonuniformity of density is due to the folding pattern of the cortex. Since the sulci on the outer surface and the gyri on the inner surface are highly folded, these regions should have more gray matter within the sphere of fixed radius as illustrated in Fig. 13. On the other hand, thin cortical regions will fold more than thick cortical regions. This inverse geometric relation is causing the negative correlation between density and thickness and, in turn, the resulting SPM differ in the regions of statistically significant difference. We further computed the correlation between these two measures across 24 subjects. Fig. 11 shows a similar result showing negative correlation across subjects in most regions of cortex.

V. CONCLUSION AND DISCUSSION

In this paper, we have presented a unified theoretical framework for WFS and the detailed numerical implementation issues. WFS is used as a smooth global parametrization of cortical surfaces. It is a very flexible functional estimation technique for scalar and vector data projected onto a unit sphere. WFS is shown to be a solution of a Cauchy problem in PDE, and for a specific weights, it becomes diffusion smoothing [10]. As a special case of WFS when the bandwidth vanishes, the traditional SPHARM can be incorporated into this more general framework. However, WFS was shown to perform better than SPHARM when data are more noisy and discontinuous by not having the significant ringing artifacts.

As an application of this novel approach, we used WFS as a tool for comparing the gray matter and the cortical thickness in a single mathematical framework. Using the WFS representation as the ground truth, cortical thickness, and gray matter density are constructed and compared. In the cortical thickness analysis, the thickness is defined using the WFS-correspondence. Afterwards, the SPM of thickness and gray matter density are compared to show the statistically significant regions do not overlap. This surprising result is caused by the negative correlation between density and thickness. Increased folding increases the gray matter density while decreasing thickness. This should serve as a spring board for more thorough investigation on comparing cortical thickness and density based morphometric techniques such as VBM.

ACKNOWLEDGMENT

The authors would like to thank S. Robbins of the Montreal Neurological Institute, J. Ashburner of the Wellcome Department of Imaging Neuroscience, M. A. Syner of the Department of Computer Science, University of North Carolina, Chapel Hill, and S. Wang of the Department of Statistics, University of Wisconsin-Madison for valuable discussions and comments.

REFERENCES

- [1] A. Andrade, F. Kherif, J. Mangin, K. J. Worsley, A. Paradis, O. Simon, S. Dehaene, D. Le Bihan, and J.-B. Poline, "Detection of fmri activation using cortical surface mapping," *Human Brain Map.*, vol. 12, pp. 79–93, 2001.
- [2] J. Ashburner and K. Friston, "Voxel-based morphometry—The methods," *NeuroImage*, vol. 11, pp. 805–821, 2000.
- [3] F. L. Bookstein, "Voxel-based morphometry should not be used with imperfectly registered images," *NeuroImage*, vol. 14, pp. 1454–1462, 2001.
- [4] C. Brechbuhler, G. Gerig, and O. Kubler, "Parametrization of closed surfaces for 3-D shape description," *Computer Vision Image Understanding*, vol. 61, pp. 154–170, 1995.
- [5] T. Bulow, "Spherical diffusion for 3-D surface smoothing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 12, pp. 1650–1654, Dec. 2004.
- [6] A. Cachia, J.-F. Mangin, D. Rivière, D. Papadopoulos-Orfanos, F. Kherif, I. Bloch, and J. Régis, "A generic framework for parcellation of the cortical surface into gyri using geodesic voronoi diagrams," *Image Anal.*, vol. 7, pp. 403–416, 2003.
- [7] M. K. Chung, "Heat kernel smoothing on unit sphere," in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2006, pp. 992–995.
- [8] M. K. Chung, K. M. Dalton, A. L. Alexander, and R. J. Davidson, "Less white matter concentration in autism: 2-D voxel-based morphometry," *NeuroImage*, vol. 23, pp. 242–251, 2004.
- [9] M. K. Chung, S. Robbins, R. J. Davidson, A. L. Alexander, K. M. Dalton, and A. C. Evans, "Cortical thickness analysis in autism with heat kernel smoothing," *NeuroImage*, vol. 25, pp. 1256–1265, 2005.
- [10] M. K. Chung, K. J. Worsley, S. Robbins, T. Paus, J. N. Taylor, J. Giedd, J. L. Rapoport, and A. C. Evans, "Deformation-based surface morphometry applied to gray matter deformation," *NeuroImage*, vol. 18, pp. 198–213, 2003.
- [11] D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans, "Automatic 3-D intersubject registration of mr volumetric data in standardized talairach space," *J. Comput. Assisted Tomogr.*, vol. 18, pp. 192–205, 1994.
- [12] R. Courant and D. Hilbert, *Methods of Mathematical Physics*. New York: Interscience, 1953, vol. 1.
- [13] A. M. Dale and B. Fischl, "Cortical surface-based analysis i. Segmentation and surface reconstruction," *NeuroImage*, vol. 9, pp. 179–194, 1999.
- [14] K. M. Dalton, B. M. Nacewicz, T. Johnstone, H. S. Schaefer, M. A. Gernsbacher, H. H. Goldsmith, A. L. Alexander, and R. J. Davidson, "Gaze fixation and the neural circuitry of face processing in autism," *Nature Neuroscience*, vol. 8, pp. 519–526, 2005.
- [15] C. Davatzikos and R. N. Bryan, "Using a deformable surface model to obtain a shape representation of the cortex," in *Proc. IEEE Int. Conf. Computer Vision*, 1995, vol. 9, pp. 2122–2127.
- [16] C. Davatzikos, A. Genc, D. Xu, and S. M. Resnick, "Voxel-based morphometry using the ravens maps: Methods and validation using simulated longitudinal atrophy," *NeuroImage*, vol. 14, pp. 1361–1369, 2001.
- [17] J. Fan and I. Gijbels, *Local Polynomial Modelling and Its Applications*. Boca Raton, FL: Chapman & Hall/CRC, 1996.
- [18] B. Fischl and A. M. Dale, "Measuring the thickness of the human cerebral cortex from magnetic resonance images," in *PNAS*, 2000, vol. 97, pp. 11050–11055.
- [19] B. Fischl, M. I. Sereno, R. Tootell, and A. M. Dale, "High-resolution intersubject averaging and a coordinate system for the cortical surface," *Hum. Brain Mapp.*, vol. 8, pp. 272–284, 1999.
- [20] J. H. Friedman, J. L. Bentley, and R. A. Finkel, "An algorithm for finding best matches in logarithmic expected time," *ACM Trans. Math. Software*, vol. 3, pp. 209–226, 1997.
- [21] K. J. Friston, A short history of statistical parametric mapping in functional neuroimaging Wellcome Dept. Imag. Neuroscience, ION, UCL, London, U.K., Tech. Rep., 2002.
- [22] A. Gelb, "The resolution of the gibbs phenomenon for spherical harmonics," *Math. Computat.*, vol. 66, pp. 699–717, 1997.
- [23] G. Gerig, M. Styner, D. Jones, D. Weinberger, and J. Lieberman, "Shape analysis of brain ventricles using spharm," in *MMBIA*, 2001, pp. 171–178.
- [24] G. Gerig, M. Styner, and G. Szekely, "Statistical shape models for segmentation and structural analysis," in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2002, pp. 18–21.
- [25] C. D. Good, I. S. Johnsrude, J. Ashburner, R. N. A. Henson, K. J. Friston, and R. S. J. Frackowiak, "A voxel-based morphometric study of ageing in 465 normal adult human brains," *NeuroImage*, vol. 14, pp. 21–36, 2001.
- [26] H. Groemer, *Geometric Applications of Fourier Series and Spherical Harmonics*. Cambridge, U.K.: Cambridge Univ. Press, 1996.
- [27] X. Gu, Y. L. Wang, T. F. Chan, T. M. Thompson, and S. T. Yau, "Genus zero surface conformal mapping and its application to brain surface mapping," *IEEE Trans. Med. Imag.*, vol. 23, no. 8, pp. 1–10, Aug. 2004.
- [28] S. C. Johnson, L. C. Baxter, L. Susskind-Wilder, D. J. Connor, M. N. Sabbagh, and R. J. Caselli, "Hippocampal adaptation to face repetition in healthy elderly and mild cognitive impairment," *Neuropsychologia*, vol. 42, pp. 980–989, 2004.
- [29] S. C. Joshi, J. Wang, M. I. Miller, D. C. Van Essen, and U. Grenander, "On the differential geometry of the cortical surface," *Vision Geometry IV*, pp. 304–311, 1995.
- [30] A. Kelemen, G. Szekely, and G. Gerig, "Elastic model-based segmentation of 3-d neuroradiological data sets," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 828–839, Oct. 1999.
- [31] K. Kollakian, "Performance analysis of automatic techniques for tissue classification in magnetic resonance images of the human brain," M.S. thesis, Concordia Univ., Montreal, QC, Canada, 1996.
- [32] J. D. MacDonald, N. Kabani, D. Avis, and A. C. Evans, "Automated 3-d extraction of inner and outer surfaces of cerebral cortex from mri," *NeuroImage*, vol. 12, pp. 340–356, 2000.
- [33] A. B. McMillan, B. P. Hermann, S. C. Johnson, R. R. Hansen, M. Seidenberg, and M. E. Meyerand, "Voxel-based morphometry of unilateral temporal lobe epilepsy reveals abnormalities in cerebral white matter," *NeuroImage*, vol. 23, pp. 167–174, 2004.
- [34] M. I. Miller, A. Banerjee, G. E. Christensen, S. C. Joshi, N. Khaneja, U. Grenander, and L. Matejic, "Statistical methods in computational anatomy," *Stat. Methods Med. Res.*, vol. 6, pp. 267–299, 1997.
- [35] M. I. Miller, A. B. Massie, J. T. Ratnanather, K. N. Botteron, and J. G. Csernansky, "Bayesian construction of geometrically based cortical thickness metrics," *NeuroImage*, vol. 12, pp. 676–687, 2000.
- [36] H.-G. Muller, "Functional modeling and classification of longitudinal data," *Scand. J. Stat.*, vol. 32, pp. 223–240, 2005.
- [37] T. Paus, A. Zijdenbos, K. J. Worsley, D. L. Collins, J. Blumenthal, J. N. Giedd, J. L. Rapoport, and A. C. Evans, "Structural maturation of neural pathways in children and adolescents: In vivo study," *Science*, vol. 283, pp. 1908–1911, 1999.
- [38] D. A. Pizzagalli, T. R. Oakes, A. S. Fox, M. K. Chung, C. L. Larson, H. C. Abercrombie, S. M. Schaefer, R. M. Benca, and R. J. Davidson, "Functional but not structural subgenual prefrontal cortex abnormalities in melancholia," *Molecular Psychiatry*, vol. 9, pp. 393–405, 2004.
- [39] A. Qiu, D. Bitouk, and M. I. Miller, "Smooth functional and structural maps on the neocortex via orthonormal bases of the laplace-beltrami operator," *IEEE Trans. Medical Imaging*, 2006.
- [40] M. Quicken, C. Brechbuhler, J. Hug, H. Blattmann, and G. Szekely, "Parameterization of closed surfaces for parametric surface description," in *IEEE Computer Soc. Conf. Computer Vision Pattern Recognit. (CVPR)*, 2000, pp. 354–360.
- [41] J. O. Ramsay and B. W. Silverman, *Functional Data Analysis*. New York: Springer, 1997.
- [42] S. M. Robbins, "Anatomical standardization of the human brain in euclidean 3-space and on the cortical 2-manifold," Ph.D. thesis, School Computer Sci., McGill Univ., Montreal, QC, Canada, 2003.
- [43] L. Shen and M. K. Chung, "Large-scale modeling of parametric surfaces using spherical harmonics," in *3rd Int. Symp. 3-D Data Processing, Visualization Transmission (3DPVT)*, Chapel Hill, NC, Jun. 14–16, 2006.
- [44] L. Shen, J. Ford, F. Makedon, and A. Saykin, "Surface-based approach for classification of 3-D neuroanatomical structures," *Intell. Data Anal.*, vol. 8, pp. 519–542, 2004.
- [45] D. O. Siegmund and K. J. Worsley, "Testing for a signal with unknown location and scale in a stationary gaussian random field," *Ann. Stat.*, vol. 23, pp. 608–639, 1996.

- [46] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in mri data," *IEEE Trans. Med. Imag.*, vol. 17, no. 1, pp. 87–97, Feb. 1988.
- [47] M. Styner, I. Oguz, S. Xu, C. Brechbuhler, D. Pantazis, J. Levitt, M. Shenton, and G. Gerig, "Framework for the statistical shape analysis of brain structures using spharm-pdm," in *Insight J.*, 2006, pp. 1–20.
- [48] P. M. Thompson, K. M. Hayashi, G. de Zubicaray, A. L. Janke, S. E. Rose, J. Semple, D. Herman, M. S. Hong, S. S. Dittmer, D. M. Dreddell, and A. W. Toga, "Dynamics of gray matter loss in Alzheimer's disease," *J. Neurosci.*, vol. 23, pp. 994–1005, 2003.
- [49] P. M. Thompson and A. W. Toga, "A surface-based technique for warping 3-dimensional images of the brain," *IEEE Trans. Med. Imag.*, vol. 15, no. 4, pp. 402–417, Aug. 1996.
- [50] G. Wahba, *Spline Models for Observational Data*. Philadelphia, PA: SIAM, 1990.
- [51] K. J. Worsley, "Local maxima and the expected euler characteristic of excursion sets of χ^2 , f and t fields," *Adv. Appl. Probab.*, vol. 26, pp. 13–42, 1994.
- [52] K. J. Worsley, S. Marrett, P. Neelin, A. C. Vandal, K. J. Friston, and A. C. Evans, "A unified statistical approach for determining significant signals in images of cerebral activation," *Human Brain Mapp.*, vol. 4, pp. 58–73, 1996.
- [53] K. J. Worsley, J. E. Taylor, F. Tomaiuolo, and J. Lerch, "Unified univariate and multivariate random field theory," *NeuroImage*, vol. 23, pp. 189–195, 2005.