# Cortical Surface Shape Analysis Based on Spherical Wavelets

Peng Yu, *Student Member, IEEE*, P. Ellen Grant, Yuan Qi, Xiao Han, Florent Ségonne, Rudolph Pienaar, Evelina Busa, Jenni Pacheco, Nikos Makris, Randy L. Buckner, Polina Golland, and Bruce Fischl*

*Abstract*—*In vivo* quantification of neuroanatomical shape variations is possible due to recent advances in medical imaging and has proven useful in the study of neuropathology and neurodevelopment. In this paper, we apply a spherical wavelet transformation to extract shape features of cortical surfaces reconstructed from magnetic resonance images (MRIs) of a set of subjects. The spherical wavelet transformation can characterize the underlying functions in a local fashion in both space and frequency, in contrast to spherical harmonics that have a global basis set. We perform principal component analysis (PCA) on these wavelet shape features to study patterns of shape variation within normal population from coarse to fine resolution. In addition, we study the development of cortical folding in newborns using the Gompertz model in the wavelet domain, which allows us to characterize the order of development of large-scale and finer folding patterns independently. Given a limited amount of training data, we use a regularization framework to estimate the parameters of the Gompertz model to improve the prediction performance on new data. We develop an efficient method to estimate this regularized Gompertz model based on the Broyden–Fletcher–Goldfarb–Shannon (BFGS) approximation. Promising results are presented using both PCA and the folding development model in the wavelet domain. The cortical folding development model provides quantitative anatomic information regarding macroscopic cortical folding development and may be of potential use as a biomarker for early diagnosis of neurologic deficits in newborns.

*Index Terms*—Folding, MRI, multiscale, neurodevelopment.

## I. INTRODUCTION

EVIDENCE suggests that morphological changes of neuroanatomical structures may reflect abnormalities in neurodevelopment, or a variety of disorders, such as schizophrenia and Alzheimer's disease (AD). These morphological variations can be characterized by the change of volume, thickness, surface area and shape. Efforts were originally made to verify the relationship between the pathology and the volumetric variation of various neuroanatomical subjects, such as the cerebral cortex, hippocampus and corpus callosum. Recently, a considerable amount of efforts has been focused on developing a technique to quantify the changes in the 2-D or 3-D shape of brain structures, which could potentially lead to more accurate diagnoses, better treatments, and an improved understanding of neurodevelopment.

To accurately study intersubject shape variations, one would like to find not only an effective shape representation but also a registration method to preserve individual variation while aligning anatomically important structures. Different techniques employed in these two aspects confer merits and disadvantages to various shape analysis methods. One of the earliest techniques developed in this field represents shapes by points sampled on the boundary of the object being studied, and the coordinates of the corresponding points on different subjects are directly used as shape features [1], [2]. Cootes *et al.* extended this method by building the point distribution model, which allows for global scale analysis of shape variation by applying principal component analysis (PCA) to the positions of the boundary points [3]. However, this method depends heavily on the accuracy of the intersubject registration for group comparison. Subsequently, parametric models were developed to decompose the boundary or surface using Fourier or spherical harmonic descriptors, and to use the decomposition coefficients as a shape descriptor [4]–[7]. A drawback of these models is the lack of ability to concisely represent local shape variation because of the global support of the basis functions. Recently, a weighted spherical harmonic representation has been developed and can be potentially used to conduct local analysis by differential weighting of the SPHARM coefficients

P. Yu is with the Harvard-MIT Division of Health Sciences and Technology (HST), Massachusetts Institute of Technology (MIT), Cambridge, MA 02139 USA (e-mail: pengyu@mit.edu).

P. E. Grant is Chief of Pediatric Radiology, Massachusetts General Hospital (MGH), Boston, MA 02114 USA and with the Athinoula A. Martinos Center for Biomedical Imaging, MGH/MIT/HMS, Charlestown, MA 02129 USA (e-mail: ellen@nmr.mgh.harvard.edu).

Y. Qi is with the Computer Science and Artificial Intelligence Laboratory (CSAIL), MIT, Cambridge, MA 02139 USA (e-mail: alanqi@csail.mit.edu).

X. Han was with the Athinoula A. Martinos Center for Biomedical Imaging, MGH/MIT/HMS, Charlestown, MA 02129 USA. He is now with the CMS Inc., St. Louis, MO 63132 USA (e-mail: xiao.han@cmsrtp.com).

F. Ségonne was with CSAIL, MIT, Cambridge, MA 02139 USA. He is now with CERTIS Laboratory, ENPC, Paris, France (e-mail: segonne@csail.mit.edu).

R. Pienaar, E. Busa, J. Pacheco and N. Makris are with the Athinoula A. Martinos Center for Biomedical Imaging, MGH/MIT/HMS, Charlestown, MA 02129 USA (e-mail: rudolph@nmr.mgh.harvard.edu; evelina@nmr.mgh.harvard.edu; jpacheco@nmr.mgh.harvard.edu; nikos@nmr.mgh.harvard.edu).

R. L. Buckner is with the Department of Psychology, Harvard University, Cambridge, MA 02138 USA, the Department of Radiology, Harvard Medical School, Boston, MA 02115 USA, and the Athinoula A. Martinos Center for Biomedical Imaging, MGH/MIT/HMS, Charlestown, MA 02129 USA (e-mail: buckner@nmr.mgh.harvard.edu).

P. Golland is with CSAIL, MIT, Cambridge, MA 02139 USA (e-mail: polina@csail.mit.edu).

*B. Fischl is with the Harvard Medical School, Boston, MA 02115 USA.
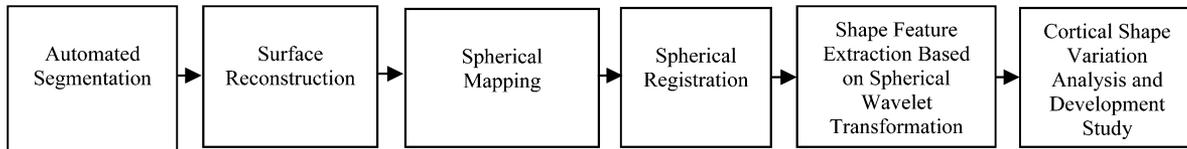
Fig. 1.   The automated procedure for conducting shape analysis of neuroanatomical structures.

[8]. Another popular method warps a template to individual subjects and studies the deformation field for shape variations [9]–[12]. Although this method is sensitive to the template selection and presents challenges in interpreting and comparing shape differences using the high-dimensional deformation field, a number of interesting shape analysis results have been obtained and more advanced techniques based on this method have been developed. Medial axis techniques, originally proposed by Blum [13] and used by Pizer *et al.* [14] and Golland *et al.* [15] in 3-D and 2-D, respectively, is a powerful tool for the shape analysis of a variety of subcortical structures. This technique allows for the separate study of the local position and thickness of the object at both coarse and fine levels. Another advantage of medial descriptions is due to an object intrinsic coordinate system, which facilitates the construction of correspondences between subjects and the subsequent statistical analysis. However, a fundamental problem of any skeletonization technique is sensitivity to perturbations in the boundary, which presents a challenge to the further development and application of medial representation in shape analysis.

The difficulties in finding both an accurate shape presentation and a robust registration method present a challenge to the study of the complex shape of the cerebral cortex in human beings, which is highly convoluted and greatly affected by neurodevelopment and neuropathy. In order to accurately and efficiently extract shape features and conduct statistical analysis, we develop a procedure to register and normalize cortical surfaces, and decompose them using spherical wavelets. The computed wavelet coefficients are used as shape features to study the folding pattern of cortical surfaces at different spatial scales and locations, as the underlying wavelet basis functions have local supports in both space and frequency.

Using this method, we studied the patterns of cortical shape variation at different spatial-frequency levels by applying PCA in the wavelet domain. A similar method was first proposed in [16] to build hierarchical active shape models of 2-D objects (such as the corpus callosum) using 1-D wavelets, which were then used for shape based image segmentation. It was further extended to learn a shape prior of 3-D objects (such as prostate and caudate nucleus) by applying PCA to the clusters of correlated spherical wavelet coefficients [17], and to use this prior for image segmentation based on the spherical wavelets presentation [18]. In this work, we emphasize on the use of PCA to study and visualize the major patterns of shape variation of cortical surfaces and the correlation of these shape variations with age and neuropsych measurements at different spatial scales by using spherical wavelets.

To model the cortical folding development of cortical surfaces from infancy to early adolescence, we fit a growth model, more exactly, the Gompertz function, to the spherical wavelet coefficients. Given a limited amount of training data, which includes a set of given MRI scans of newborns and children, we employ a regularization framework to improve prediction performance on new MRI scans. We develop an efficient method to estimate this regularized Gompertz model based on Broyden–Fletcher–Goldfarb–Shannon (BFGS) approximation [19].

The entire procedure, including MR image preprocessing, spherical wavelet transformation, statistical analysis using PCA, and the cortical folding development model are introduced in detail in Section II. Although the entire procedure can be used to analyze both the gray matter (GM)/white matter (WM) boundary and GM/cerebrospinal fluid (CSF) boundaries, only the gray/white surface, which is a direct reflection of the gyral folding, is used in this paper to exemplify the developed methods. The results of using PCA in detecting the multiresolutional patterns of shape variation in a nondemented aged population are demonstrated in Section III. We also present the use of the proposed folding development model in detecting spatial scales and patterns of the cortical folding development of the GM/WM boundary in newborns and children.

## II. METHODS

The automated procedure for conducting shape analysis using the spherical wavelets is shown in Fig. 1. The details of each step are introduced in this section. The tools used for preprocessing the cortical surfaces, and the procedures developed to transform the reconstructed cortical surfaces using SPHARM and spherical wavelets are first introduced. Then we describe the procedure developed to study the pattern of shape variations in a population based on the PCA technique. At last, the regularized Gompertz model that is used to study the folding development of cortical surface is presented.

### A. Preprocessing

To decompose a cortical surface using basis functions defined in a spherical coordinate system, such as spherical wavelets, the surface is first mapped onto a parameterized sphere. We then use a registration procedure to establish the correspondence across subjects in order to carry out the statistical analysis. A set of automated tools distributed as part of the FreeSurfer package[1] are used to preprocess the data, which includes cortical surface reconstruction, spherical transformation, and spherical registration based on the folding patterns of cortical surfaces.

To reconstruct the cortical surfaces, which include the GM/WM boundaries (hereinafter referred as WM surfaces) and gray matter (GM)/CSF boundaries (hereinafter referred as pial surfaces) of the left and right hemispheres, the MR images are first registered to a prebuilt template in the Talairach space. The image intensity is normalized to remove spatial variations induced by inhomogeneities in the RF field, and used to guide

[1]http://www.surfer.nmr.mgh.harvard.edu.
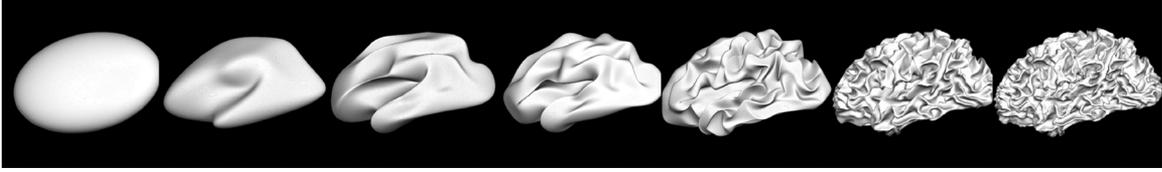
Fig. 2. Reconstructed cortical surfaces using SPHARM coefficients truncated at degree 1, 2, 5, 10, 20, and 40 and the original surface (from left to right). This figure demonstrates SPHARM's effectiveness for multi-resoluational representation of cortical surfaces. Note that by using a higher degree, one can always represent the surface more accurately. However, since SPHARM bases are global, SPHARM coefficients cannot represent local shape changes concisely.

skull stripping and WM labeling. This WM segmentation is further refined and cut to generate a single connected mass of each hemisphere. A surface tessellation is then constructed for each WM volume by representing each square face separating the WM voxels from other classes with two triangles as detailed in [20]. To generate a more accurate and smoother hemisphere, this tessellation is refined and deformed in the normalized image volume under smoothness and boundary intensity constraints. Furthermore, the WM surface is deformed outwards to the location in the volume that has the largest intensity contrast between the GM and CSF, and refined to generate the pial surface. Finally, topological defects are automatically detected and corrected for both surfaces to guarantee spherical topology.

Next, the reconstructed WM surface is mapped onto a sphere in two steps. First, the WM surface is inflated and projected radially to a sphere. Then, the large folds and metric distortion introduced by the projection process are removed by the minimization of folded area and the preservation of the local and long range distances between vertices [21]. With this spherical representation, surfaces of different subjects are then registered in the spherical coordinate system by minimizing an energy functional that is a combination of a topology preserving term, a folding alignment term and a metric preservation term. The alignment of gyral and sucal patterns enables us to find anatomically corresponding points on the reconstructed cortical surfaces across subjects, and the metric preservation (including area and distance) allows the preservation of individual variations. The effectiveness of this spherical registration technique in finding correspondence across subjects is proven in [22].

### B. Spherical Harmonics (SPHARM)

The common spherical coordinate system established by this procedure allows us to extract shape features using SPHARM and spherical wavelets. As a natural extension of Fourier transformation on the sphere, SPHARM has been demonstrated to be a powerful tool in describing the boundary of objects of spherical topology. In this method, the coordinates $\overline{v}(\theta, \varphi) = (x, y, z)$, $\theta \in [0, \pi)$, $\phi \in [0, 2\pi)$, of a parameterized surface are expressed as the weighted summation of a set of spherical harmonic basis functions of degree $l$ and order $m$, $Y_l^m$, $-l \leq m \leq l$

$$\overline{v}(\theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \overline{c}_l^m Y_l^m(\theta, \varphi) \tag{1}$$

where the coefficients $\overline{c}_l^m$ contain hierarchical shape information. Truncating the spherical harmonic series at different degrees results in object representations at different levels of details, as shown in Fig. 2. However, SPHARM coefficients

cannot concisely indicate the location of shape differences because of the global support of the basis functions. Most recently, a weighted SPHARM technique has been developed and applied to cortical thickness analysis and voxel-based morphometry [8]. This new technique provides a framework for weighting the SPHARM coefficients differently and can be potentially used to improve the localization property of SPHARM representation.

In this paper, both SPHARM and wavelets methods are implemented to compare their abilities in capturing local shape variations as shown in Section III. SPHARM coefficients are estimated by solving for the least square solution of the linear equation (1). Therefore, the SPHARM coefficients estimation (highest order $= 40$) of a typical cortical surface (12 000 vertices) constructed using methods detailed above requires about one hour of computation time on a standard PC architecture. The spherical wavelet transformation is more efficient (about one minute for typical cortical surface) because it has a computational time linear to the number of vertices. However, computational cost of SPHARM can be greatly minimized with advanced numerical implementation, such as the iterative residual fitting method proposed in [23].

### C. Spherical Wavelets

Broadly speaking, a wavelet representation of a function consists of a coarse overall approximation together with detail coefficients that influence the function at various spatial scales and locations. The classical form of wavelet analysis decomposes signals onto a set of basis functions, called wavelets, in which every wavelet is a scaled and translated copy of a single unique function, called the mother wavelet [24]. However, this shift-invariant theory breaks down when representing functions defined on a bounded surface, such as a sphere. No longer relying on the adaptive constructions based on traditional dilation and shifting, the spherical wavelets employed in this work belong to the *second generation wavelets*, which maintain the notion that a basis function can be written as a linear combination of basis functions at a finer, more subdivided level.

The construction of these spherical wavelets relies on a recursive subdivision of an icosahedron (subdivision level 0) (e.g., [25]). Denoting the set of all vertices on the mesh before the $j$th subdivision as $K(j)$, a set of new vertices $M(j)$ can be obtained by adding vertices at the midpoint of edges and connecting them with geodesics. Therefore, the complete set of vertices at the $(j + 1)$th level is given by $K(j + 1) = K(j) \cup M(j)$. As a result, the number of vertices at level $j$ is $10 \times 4^j + 2$, e.g., 12 vertices at level 0, 42 at level 1, 162 at level 2, etc. Next, an interpolating subdivision scheme is used to construct the scaling

functions $\varphi_{j,k}$ defined at level $j$ and node $k \in K(j)$ such that a scaling function at level $j$ is a linear combination of the scaling functions at a finer level $j+1$. Using these scaling functions, the wavelet $\psi_{j,m}$ at level $j$ and node $m \in M(j)$ can be constructed by the lifting scheme, the basic idea of which is to start with a simple construction of wavelet, and then update it to a new, more complete one. Specifically, we first define the wavelet function as the scaling function at a higher level, and then "lift" it so that it has a vanishing integral. Using the scaling function at level 0 and wavelets at level 0 and higher, a basis for the function space $L_2 = L_2(S^2, d\omega)$, where $d\omega$ is the usual area measure, is then constructed so that any functions with finite energy can be decomposed as a linear combination of these basis functions.

Spherical wavelets constructed in this way have local support in both space and frequency. As detailed in Appendix A, the transformation is easy to carry out without explicit construction of the wavelet and scaling functions and has a computation time linear to the vertices number [25]. Note that these spherical wavelets are only biorthogonal [26] (wavelets at the same level and between different levels are not orthogonal to each other such that they are correlated) because currently there are no wavelet bases on the sphere that consist of functions that are orthogonal, compactly supported, symmetric, and smooth [27]. To assess the correlation between a pair of wavelet basis functions, we calculate their correlation coefficients as

$$\frac{\int \psi_{jm}\psi_{j'm'}d\omega}{\sqrt{\int |\psi_{jm}|^2 d\omega \int |\psi_{j'm'}|^2 d\omega}} \tag{2}$$

where $j = j'$, $m \in Neighbor(m')$ for a pair of wavelet functions at the same level, and $j = j' + 1$, $m \in M(j)$, $m' \in M(j')$, $m' \in Neighbor(m)$ for a pair of wavelet functions across the consecutive levels. The averaged correlation coefficients over all the pairs are 0.0225 at the same level and 0.106 across the adjacent levels, indicating fairly weak correlations between wavelet basis functions. Therefore, an approximately orthogonal decomposition can be expected using this spherical wavelet bases. Note that it is possible to apply PCA to completely orthogonalize the wavelet functions. However, the orthogonal PCA bases do not enjoy the locality property of biorthogonal wavelets in the spatial-frequency domain and this locality property is key to our subsequent analyses.

This biorthogonal spherical wavelets are used to transform cortical surfaces reconstructed using the previously described procedure. Because the reconstructed cortical surfaces are mapped onto a sphere and registered in a spherical coordinate system, the original spatial coordinates of points on a cortical surface can be considered as functions defined on the sphere. To transform them into the wavelets domain, the spatial coordinates of each subject's cortical surface are first interpolated onto the mesh of an icosahedron (subdivision level $= 7$) based on their corresponding spherical coordinates established by the spherical registration. We choose an icosahedron at subdivision level 7 to represent a parameterized sphere because it has a total number of 163 842 vertices and is thus sufficient to represent the spherical map of a cortical surface reconstructed

from $\sim$1 mm isotropic MRI, which typically has about 120 000 vertices. An icosahedron subdivided 6 levels has only 40 962 vertices and may lead to loss of useful surface details if it is used to represent the finest level sphere parameterization. Conversely, further increasing the resolution of the spherical parameterization will unnecessarily and significantly increase the computation time (an icosahedron at subdivision level 8 has 655 362 vertices).

Unlike other intrinsic shape features such as curvatures, spatial coordinates of a surface depend on a rather arbitrary coordinate system that each individual subject is originally represented in. In order to make the coordinate functions invariant to rotation, translation and scaling and thus to be used as valid shape features, they have to be normalized with respect to a common reference coordinate frame. This normalization is initialized by first transforming each spatial coordinate function using the transformation matrix calculated previously for volume registration during surface reconstruction. Then the roughly normalized coordinate functions of all the surfaces are averaged to create a new template surface in the spatial domain for the second round normalization. Finally, each coordinate function is normalized by finding an optimal affine transformation that minimizes the mean square error between the transformed individual surface and the template. The normalization process simply aims to bring each subject into the same coordinate system and remove the arbitrary affine components in their coordinate functions; intersubject shape variations are still preserved after the normalization. In addition, since the surface correspondence is already found by the spherical registration, the affine normalization is robust to compute.

The normalized coordinate vector $\overline{v} = (x, y, z)^T$ is then expanded by the scaling function at the ground level and a set of spherical wavelet functions as

$$\overline{v} = \sum_{k=K(0)} \overline{\lambda}_{0,k}\varphi_{0,k} + \sum_{j=0,\dots6, m\in M(j)} \overline{\gamma}_{j,m}\psi_{j,m} \tag{3}$$

where $\overline{\gamma}_{j,m}$ is the 3-D wavelet coefficient corresponding to the $x$, $y$, and $z$ coordinates at level $j$, $j = 0, \dots 6$ and location $m$, $m \in M(j)$. To facilitate notation we define for the coarsest level $M(-1) = K(0)$, $\psi_{-1,k} = \varphi_{0,k}$ and $\overline{\gamma}_{-1,m} = \overline{\lambda}_{0,k}$ so that the wavelet coefficients at level $-1$ is actually the scaling coefficients at the ground level.

Fig. 3 illustrates the decomposition of a cortical surface starting from the finest level. At each level, the cortical surface is split into a low-resolution part and a detail part (wavelet coefficients). As shown in the rightmost figure in Fig. 3, there are 12 wavelet coefficients at level $-1$, each of which represents the overall shape of the cortical surface in the region around a vertex on the icosahedron. At subsequent levels, wavelet coefficients provide descriptions of the spatial variations of the surface at increasingly finer resolutions. Moreover, the larger a wavelet coefficient is, the deeper the corresponding surface folding is at that specific location and resolution.

In the following sections, we use wavelet decompositions to study the shape of cortical surfaces locally and hierarchically.
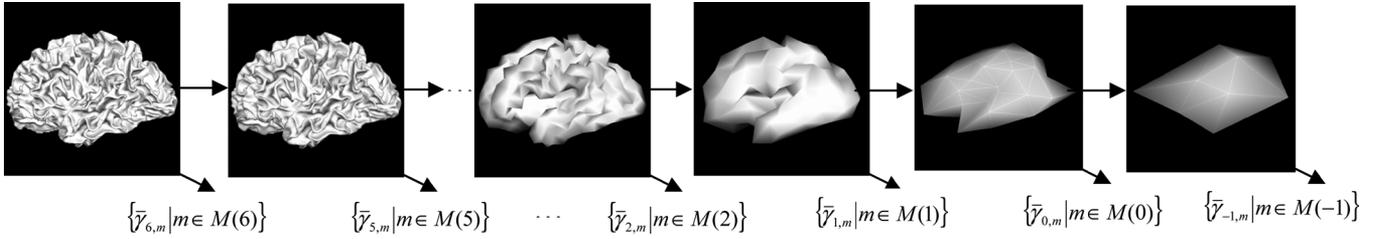
Fig. 3.　Spherical wavelet decomposition of cortical surface from level 7 (the leftmost image). At each level, the surface is further decomposed into lower resolution approximation and wavelet coefficients. Wavelet coefficients at level −1 quantify the overall shape of the cortical surface. Wavelet coefficients at subsequent levels encode the surface spatial variations at finer resolutions.

## D. Principal Component Analysis (PCA)

PCA is a useful tool for finding patterns in high-dimensional data and has been extensively used in computer vision and image recognition. For example, PCA has been applied to build generative models of shape variations within a single population and to segment 2-D or 3-D medical images [28]–[30]. The basic idea of these PCA-based approaches is to identify and visualize the first few principal modes of variation of the geometry of a particular structure across a group of subjects.

For example, let $x_i(i = 1, \cdots N)$ denote a set of vectors containing shape features such as the coordinates of surface points for a group of subjects, then any individual shape can be decomposed as

$$x_i = \overline{x} + \sum_{n=1}^{N} e_n e_n^T (x_i - \overline{x}) \qquad (4)$$

where $\overline{x}$ is the mean of $x_i(i = 1, \cdots, N)$, and $e_1, \cdots, e_N$ are the eigenvectors corresponding to eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_N$ of the covariance matrix of $x$, in decreasing order. This is essentially equivalent to linearly transforming a dataset into a new coordinate system such that the variance of the projection of the dataset on the first axis is the greatest, and the variance of projection on the second axis is the second greatest, and so on. The fact that the variance explained by each eigenvector is equal to the corresponding eigenvalue enables us to study the most significant modes of variation in the dataset. Usually, most of the variations can be sufficiently represented by a small number of modes, $k$, so that the sum of the first $k$ variances represents a sufficiently large proportion of total variance of all the variables used to derive the covariance matrix. By limiting the number of terms in (4), PCA can decrease the data dimension and remove data noise in the subspace spanned by the eigenvectors corresponding to relatively small shape variations. However, it has been argued that omitting these eigenvectors leads to the failure of characterizing subtle, yet important shape features because coordinates of all the points on the surface are collected in the shape feature vector and consequently all the eigenvectors contain useful shape information [31].

To avoid losing localized shape information of cortical surfaces, we conduct PCA on the wavelet coefficients at different frequency levels separately, as the coefficients in the lowest level provide an overall approximation and localized morphological variations are captured hierarchically by the higher-level coefficients. Instead of using coordinates of all the points on the surface as shape features in (4), each time we take only as input

$x_{ij} = \{\overline{\gamma}_{j,m}^i | m \in M(j)\}$, the subset of the wavelet coefficients at the $j$th frequency level. Once the set of eigenvectors $e_{jn}(j = -1, \cdots 6, n = 1, \ldots N)$ that characterize the majority of the variance of the wavelet coefficients at the $j$th frequency level is found, the corresponding shape variations can be visualized by inversely transforming the principal components to generate the principal surfaces. This visualization technique provides an intuitive way to analyze and understand the most distinct patterns of shape variations within a group of subjects from coarse to fine resolution.

## E. Cortical Folding Development Model

The human cortex is highly convoluted, in contrast to the smooth cortex found in other animals such as mice and rats. In human beings, cortical development begins prenatally, and the majority of neurons are generated before birth. The development of cortical folding starts at about 9 weeks in gestation, changes dramatically until birth, but continues into late adolescence. The mechanism involved in the regulated formation of folding pattern remains unclear. It is hypothesized that folding pattern formation is caused by neuron differentiation, migration and the growth of neutrite. Another theory suggests that differential growth of the outer layers relative to inner layers of the cortex results in cortical buckling [32]. A third theory proposed that the mechanical tension generated during the "long-distance" connections of different regions of the brain leads to the formation of folding [33]. Like many growth phenomena in nature, the folding of the human cortex starts slowly, and accelerates before slowing down to approach a limit. In this paper, we model the folding development of the gray/white boundary at different spatial scales using a growth model, more exactly, the Gompertz model [34], in the wavelet domain. Unlike PCA, which is a linear data model, the Gompertz model is a nonlinear one that captures both the fast growth and saturation phases of cortical folding development. Specifically, if $w(t)$ is one of the spherical wavelet features extracted from a subject at age $t$, we use the Gompertz curve [34] to model the change of this feature at different ages as follows:

$$w(t_i) = g_1 \exp\left(-\exp\left(-g_2(t_i - g_3)\right)\right) + n(t_i), \quad i = 1, \cdots, N \qquad (5)$$

where $g_1$ is the maximum value at mature, $g_2$ is the growth rate that quantifies the speed of the folding development, $g_3$ is at the inflexion point and indicates the age of the fastest folding development, and $n(t)$ represents additive noise with mean zero.

Due to the limited number of subjects available in this paper, we apply a regularization framework for parameter estimation

to avoid overfitting. In such a framework, we minimize a cost function over variables $g_1$, $g_2$, and $g_3$

$$Q(g_1, g_2, g_3) = \sum_{i=1}^{N} \left( g_1 \exp\left(-\exp\left(-g_2(t_i - g_3)\right)\right) - w(t_i)\right)^2$$
$$+ c \sum_{j=1}^{3} g_j^2 \quad (6)$$

where the first term on the right-hand side of (6) models the empirical error of model fitting, and the second term is a scaled $l^2$-norm regularizer with the scaling factor $c$ controlling the trade-off between the empirical error and the degree of regularization. Having the weighted sum of the squared parameters in the cost function, in addition to the empirical mean square error, constrains the model space and avoids overfitting to data noise. This method is a special form of regularization, which is known as weight decay in statistical learning theory [35].

To minimize the cost function $Q$, we first compute its gradient, which has the closed form

$$\frac{dQ}{dg_1} = 2 \sum_{i=1}^{N} \left( g_1 \exp\left(-\exp\left(-g_2(t_i - g_3)\right)\right) - w(t_i)\right)$$
$$\times \exp\left(-\exp\left(-g_2(t_i - g_3)\right)\right) + 2cg_1$$
$$\frac{dQ}{dg_2} = 2 \sum_{i=1}^{N} \left( g_1 \exp\left(-\exp\left(-g_2(t_i - g_3)\right)\right) - w(t_i)\right)$$
$$\times \exp\left(-g_2(t_i - g_3)\right) g_1$$
$$\times \exp\left(-\exp\left(-g_2(t_i - g_3)\right)\right) (t_i - g_3) + 2cg_2$$
$$\frac{dQ}{dg_3} = -2 \sum_{i=1}^{N} \left( g_1 \exp\left(-\exp\left(-g_2(t_i - g_3)\right)\right) - w(t_i)\right)$$
$$\times \exp\left(-g_2(t_i - g_3)\right) g_1$$
$$\times \exp\left(-\exp\left(-g_2(t_i - g_3)\right)\right) (g_2) + 2cg_3. \quad (7)$$

Since a simple gradient method suffers from slow convergence, we adopt a quasi-Newton method based on the BFGS approximation [19] of the Hessian matrix. The BFGS method enables us to efficiently minimize $Q$ over the parameters $\{g_i\}_{i=1,\ldots 3}$. We tune the regularization parameter $c$ based on the leave-one-out cross-validation. Specifically, we compute the mean square error of our predictions on the held-out data point using the model parameters optimized from the rest of the training data. From a collection of prespecified values, we select the parameter $c$ that minimizes the leave-one-out error.

Furthermore, we estimate the Bayesian confidence intervals of our estimated parameters $\mathbf{g} = \{g_i\}_{i=1,\ldots 3}$. Note that using the regularized cost function (6) is equivalent to using a Bayesian model; if we scale (6) by the observation noise variance (assuming it is known), then the log likelihood function $\log p(\mathbf{w}|\mathbf{g}, \mathbf{t})$, $\mathbf{w} = \{w_i\}_{i=1,\ldots N}$, $\mathbf{g} = \{g_i\}_{i=1,\ldots 3}$, $\mathbf{t} = \{t_i\}_{i=1,\ldots N}$ in the Bayesian model corresponds to the first term on the right-hand side of (6) up to a normalization constant, and the Gaussian prior distribution $p(\mathbf{g})$ corresponds to the second term on the right-hand side of (6) up to a normalization constant. Therefore, minimizing (6) amounts to finding the maximal value of the posterior distribution $p(\mathbf{g}|\mathbf{w}, \mathbf{t})$.

Although we can efficiently compute the parameters $\widehat{\mathbf{g}}$ that maximize the posterior distribution, it is computationally intractable to calculate the exact distribution $p(\mathbf{g}|\mathbf{w}, \mathbf{t})$ because the likelihood function is non-Gaussian in $\mathbf{g}$. Therefore, we use Laplace's approximation [36] to efficiently approximate the exact posterior distribution as a Gaussian

$$p(\mathbf{g}|\mathbf{w}, \mathbf{t}) \approx p(\widehat{\mathbf{g}}|\mathbf{w}, \mathbf{t}) \exp\left(\frac{1}{2}(\mathbf{g} - \widehat{\mathbf{g}})^T \mathbf{H}(\mathbf{g} - \widehat{\mathbf{g}})\right) \quad (8)$$

where the Hessian matrix $\mathbf{H}$ is calculated by

$$\mathbf{H} = \left. \frac{d^2 \log\left(p(\mathbf{g}|\mathbf{w}, \mathbf{t})\right)}{d\mathbf{g} d\mathbf{g}^T}\right|_{\mathbf{g}=\widehat{\mathbf{g}}} = -\frac{1}{2\sigma_w^2} \left. \frac{d^2 Q(\mathbf{g})}{d\mathbf{g} d\mathbf{g}^T}\right|_{\mathbf{g}=\widehat{\mathbf{g}}}. \quad (9)$$

Once we have computed the Hessian matrix, the variance matrix of the posterior distribution $p(\mathbf{g}|\mathbf{w}, \mathbf{t})$ can be approximated as $-\mathbf{H}^{-1}$. Thus the 90% Bayesian confidence intervals of the estimated parameters are obtained by $\{\widehat{g}_i \pm 1.64\sigma_i\}_{i=1,\ldots 3}$, where $\{\sigma_i\}_{i=1,\ldots 3}$ are the diagonal components of the approximate variance matrix.

The regularized Gompertz model is applied to study the development of cortical folding in newborns based on spherical wavelets. To measure the goodness-of-fit of the model, we calculate the $R^2$ statistic, the ratio of the sum of squares explained by the model and the total sum of squares around the mean

$$R^2 = 1 - \frac{\sum_{i=1}^{N} \left(\widehat{w}(t_i) - w(t_i)\right)^2}{\sum_{i=1}^{N} \left(\widehat{w}(t_i) - \frac{1}{N} \sum_{i=1}^{N} w(t_i)\right)^2}. \quad (10)$$

We report the estimated parameters with 90% confidence interval and the $R^2$ statistics of the wavelet coefficients fitting at different frequency levels in Section III. Fig. 7 shows the fitted Gompertz curve at different frequency overlaid on the original data. These results demonstrate the effectiveness of the regularized Gompertz model for cortical folding development.

### F. Data

Two sets of high-resolution structural MR scans were analyzed in this paper. The first dataset was obtained from a total of 84 nondemented older participants (42 women: 67–95, mean age = 80, standard derivation = 7.25; 42 men: 71–94, mean age = 79, standard derivation = 7.17). These data have been reported previously in several publications associated with the Washington University Alzheimer's Disease Research Center (ADRC). None of the participants had any history of neurologic, psychiatric, or medical illness that could contribute to dementia or a serious medical condition. Two to four high-resolution MP-RAGE scans were motion corrected and averaged per participant (four volumes were averaged for all except five participants; Siemens 1.5T Vision System, resolution $1 \times 1 \times 1.25$ mm, TR $= 9.7$ ms, TE $= 4$ ms, FA $= 10°$, TI $= 20$ ms, TD $= 200$ ms) to create a single high contrast-to-noise image volume. These acquisition parameters were empirically optimized to increase GM/WM and GM/CSF contrast. Cortical surfaces were reconstructed and registered as described in the previous section. This dataset was mainly used

to study normal variations and aging-related shape changes of WM surface in a healthy older population.

The second dataset was from eight normal neonates with corrected gestational ages (cGA) of 30.57, 31.1, 34, 37.71, 38.1, 38.4, 39.72, and 40.43 weeks, and three children who were approximately 2, 3, and 7 years old at the time of scanning. T1 weighted 3-D SPGR images were collected on a 1.5T scanner, with TR/TE $= 30/8$, flip angle $= 25 - 30°$, matrix $= 256 \times 192$, FOV $= 220 \times 165$ mm or $200 \times 150$ mm and slice thickness 1.2 to 1.4 mm. The images of newborns were manually segmented into WM and cortical regions due to inverted gray/white contrast and low contrast of the GM/WM boundary. Based on the manual segmentation, the cortical surfaces are reconstructed using Freesurfer tools. The children dataset was processed and registered with newborns using the automated tools described above. Wavelet transformation and regularized Gompertz model are then applied to the reconstructed WM surfaces to study the folding development of cortical surface. To be compared with neonates, children's ages were converted to 167, 235, and 451 weeks by assuming a 40 week gestation period.

## III. RESULTS

### A. Comparison of Spherical Wavelets With SPHARM

To compare the abilities of SPHARM and spherical wavelets to detect local shape variation, both methods were applied to decompose an inflated cortical surface model with a synthesized shape deformation. The cortical surface model is made by interpolating an inflated cortical surface onto an icosahedron at subdivision level 4 (2562 vertices). The deformation is simulated by moving the first- and second-order neighbors of a vertex, denoted as $v$, outwards in their normal direction for 4 mm. After spherical wavelet transformation, the total number of wavelet coefficients is $2562 \times 3 = 7686$, because each of the $x$, $y$, and $z$ components in the coefficients vector $\overline{c}_{j,m}$ and $\overline{\gamma}_{j,m}$ [as in (1) and (3)] is evaluated individually in this comparison. The highest order of SPHARM decomposition is chosen to be 60 to achieve a comparable accuracy in shape representation, which results in a total number of $61 \times 61 \times 3 = 11163$ SPHARM coefficients.

With two sets of coefficients computed for both the original and deformed surfaces, a new set of coefficients is made by replacing a certain number of the original surface's coefficients with deformed surface's coefficients that are affected the most by this deformation, using both SPHARM and spherical wavelets methods. The coefficient variation is calculated by $\||c_d| - |c_o|\|$, where $c_d$ and $c_o$ are the corresponding coefficients calculated from the deformed and original surfaces respectively. We then reconstruct the deformed surface using this new coefficients set and measure the reconstruction error by $(1/N) \sum_{i \in \{N_1(v), N_2(v)\}} |\overline{x}_r^i - \overline{x}_d^i|$, where $\overline{x}_r^i$ and $\overline{x}_d^i$ are the coordinates of the $i$th vertex on the reconstructed and deformed surfaces respectively, $N_1(v)$ and $N_2(v)$ denote the first and second neighbors of the vertex $v$, and $N$ is the total number of vertices in the deformed region.

Fig. 4(a) is the original surface and Fig. 4(b) shows the set of wavelet coefficients with variation values larger than 0.1, with red dots indicating the centers of their support regions. Fig. 4(b)
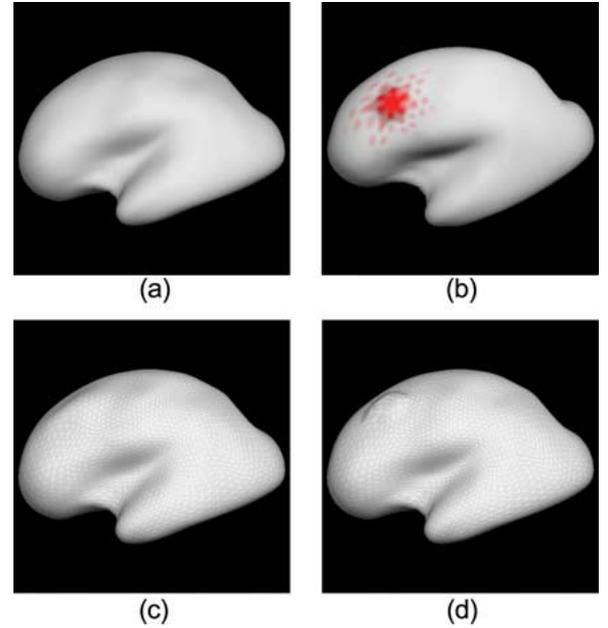


Fig. 4. The comparison of SPHARM and spherical wavelets in representing local shape variation (a) original inflated cortical surface. (b) The deformed surface with red dots indicating the location of the wavelet coefficients that are most affected by the simulated deformation. (c) Surface reconstructed with a new set of SPHARM coefficients made by replacing 100 of the original surface's coefficients with deformed surface's coefficients that are most affected by the deformation. (d) Surface reconstructed with 100 most varied spherical wavelet coefficients.

illustrates the conciseness of the wavelet representation as only the coefficients of wavelet functions in the neighborhood of the deformation region are affected. Conversely, the majority of the SPHARM coefficients have variations larger than 0.1. A quantitative comparison of the number of most varied coefficients used to have the same reconstruction error is given in Table I. With the top 100 most varied SPHARM coefficients, the reconstruction error is 2.5 mm (shown in Table I) and the synthetic bump is not well reconstructed, as shown in Fig. 4(c). On the contrary, the deformation is well recovered using 100 most varied spherical wavelet coefficients as shown in Fig. 4(d). With the use of all the 11163 SPHARM coefficients, the reconstruction error decreases to 0.0013 mm. This result verifies that SPHARM can be used to accurately represent cortical surface, but just not as concisely because local shape variation can cause changes of a large amount of coefficients. The compactness and the data compression nature of the wavelet transformation are particularly important to avoid multiple comparison problems that often plague the statistical analysis of neuroimaging data.

### B. Detection of Shape Variation Using PCA

The PCA study of the wavelet coefficients in the nondemented older population demonstrated a wide range of differences of cortical surface geometry, in both the overall shape of the cortex and the hierarchically finer local details. Most of the shape variance (98%) was represented by the first 10 to 20 eigenvectors and the variance explained by the first principal component ranges from 8% to 13% of the total variance at the lower spatial-frequency levels. Shape variances

TABLE I
COMPARISON OF THE NUMBER OF MOST VARIED COEFFICIENTS USED TO
ACHIEVE THE SAME RECONSTRUCTION ERROR (mm) USING SPHARM AND
SPHERICAL WAVELETS METHODS. SURFACE IS RECONSTRUCTED WITH A
NEW SET OF COEFFICIENTS MADE BY REPLACING A CERTAIN NUMBER
OF THE ORIGINAL SURFACE'S COEFFICIENTS WITH DEFORMED SURFACE'S
COEFFICIENTS THAT ARE MOST AFFECTED BY THE SYNTHETIC DEFORMATION.
RECONSTRUCTION ERROR IS MEASURED AS THE ROOT MEAN SQUARE
DIFFERENCE BETWEEN THE RECONSTRUCTED SURFACE AND DEFORMED
SURFACE IN THE DEFORMED REGION

| Reconstruction Error (mm) | 1 | 1.5 | 2.5 |
|---|---|---|---|
| SPHARM | 500 | 300 | 100 |
| Wavelets | 50 | 27 | 7 |

of higher spatial scales were spread out more evenly over 50 to 80 eigen-components.

The shape variation represented by the $n$th principal component at the $j$th frequency level is illustrated by generating two sets of new wavelet coefficients

$$\gamma_{j\pm} = \overline{\gamma}_j \pm 3\sigma_{nj}e_{nj} \tag{11}$$

where $\overline{c}_j$ is the mean wavelet coefficients of all the subjects at the $j$th level, and $\sigma_{nj}^2$ is the $n$th eigenvalue of the covariance matrix of the wavelet coefficients at level $j$. Then, a new coefficient set consisting of wavelet coefficients at all levels is generated by filling the other levels with the mean wavelet coefficients. At last, two synthetic surfaces $S_{\pm 3\sigma}$ are generated by inversely transforming these two sets of wavelet coefficients. The difference between the two surfaces represents the shape variations characterized by the corresponding eigen-mode at different frequency levels.

As an example, Fig. 5 shows the surfaces generated at levels $-1$ to 1 with the color indicating the $l_2$ norm, location and support region of each coefficient in the first eigenvector $e_{1j}$. Since the magnitude of wavelet coefficient quantifies the amount of shape variations, the lighter the color is, the more folded the surface is in the localized region. Note that one cannot directly compare the colors at different levels, since the wavelet functions are not normalized across levels in these results. This is because the normalization would lead to a wild difference between the wavelet coefficients across levels, masking the fine details in visualization. Because the support regions of neighbor wavelet basis functions overlap with each other, the color of each point on the cortical surface is assigned by the $l_2$ norm of the coefficient whose center of support is the closest. A set of arrows are used to point to the regions that vary the most across subjects at each level. The real surfaces (hereinafter referred as real $\pm 3\sigma$ surfaces) that have the largest positive and negative projections on the first eigen-component are also shown in Fig. 5 for comparison. The shape difference between the $S_{\pm 3\sigma}$ synthetic surfaces (first row and second row) is clearly visible in the real surfaces as well (third and four rows), which verifies the detected shape variations by the PCA method.

Finally, a preliminary study of cortical shape (GM/WM boundary) variations due to healthy aging was carried out by observing the change with age of the surfaces projected in the space spanned by the first a few principal components at each level. Specifically, the projected surface of the $i$th subject at

the $j$th level is reconstructed by inversely transforming a new set of wavelet coefficients containing the projected wavelet coefficients on the set of eigen-components that represent 98% of the total variation at the $j$th level, and the mean wavelet coefficients at other levels.

Shape changes consistent with age were observed in the low-frequency domain as well. Fig. 6 shows the projected cortical surfaces in three age ranges of female and male subjects at level 1. The colormap encodes the $l_2$ norm, location and support region of the projected coefficients of each subject on the set of eigenvectors representing 98% of the total variance. An arrow is pointed to a region of decreasing intensities with age on the precentral sulcus, indicating a decreased folding of this area. Another arrow is pointed to a region of increasing intensities with age on the occipital lobe, indicating an increased folding in this area. Visual inspection of the whole population confirms the corresponding shape variations are the narrowing of the central sulcus and the elongation of the occipital lobe, which may characterize and correlate with WM atrophy. However, these changes were not salient in male groups as the color intensities in these two regions do not vary much as shown in the second row of Fig. 6. Similarly, although some regions in the insula, orbito-frontal cortex and medial frontal lobes have high intensity values, indicating large folding in these areas, they do not vary with age.

These results were verified by regressing the projections of all the subjects on the first principal component with age, gender, and five other neuropsych measurements including Mini Mental State Examination (MMSE), Wechsler Memory Scale (WMS) Long Memory, WMS Digit Span, WMS Associates Recall Easy, and WMS Associates Recall Hard. The shape variations detected using PCA are significantly correlated with age at level $-1$, 1, 3, and 4 by regressing out all the other factors. WMS Assoc Recall Easy is significantly correlated with shape variations at level $-1$ and 0 by regressing out all the other factors. More rigorous study will be carried out to investigate the relationship between neuropsych measurements and cortical shape variations.

Although only results at the lower frequency levels are shown in Figs. 5 and 6, the visualization methods introduced can be used to examine the most distinguished shape variations at all levels. However, caution should be used in interpreting these results because the shape changes are not entirely uncorrelated within neighbor regions, as discussed in Section II.

### C. The Folding Development Study of GM/WM Boundary

In this section, we describe the results of the regularized Gompertz model to detect and characterize the development of the folds of the GM/WM boundary in the neonate and child population.

Specifically, we first fit the regularized Gompertz function [see (5)] to the mean squared values of all the wavelet coefficients at different frequency level. Since the $l^2$-norm $|\overline{\gamma}_{j,m}|$ of each wavelet coefficient quant-ifies the amount of spatial variations at a certain resolution and location, the wavelet power quantifies the degree of folding at that spatial resolution across the whole brain. Therefore, by modeling the change of the wavelet power with age using a regularized Gompertz function,
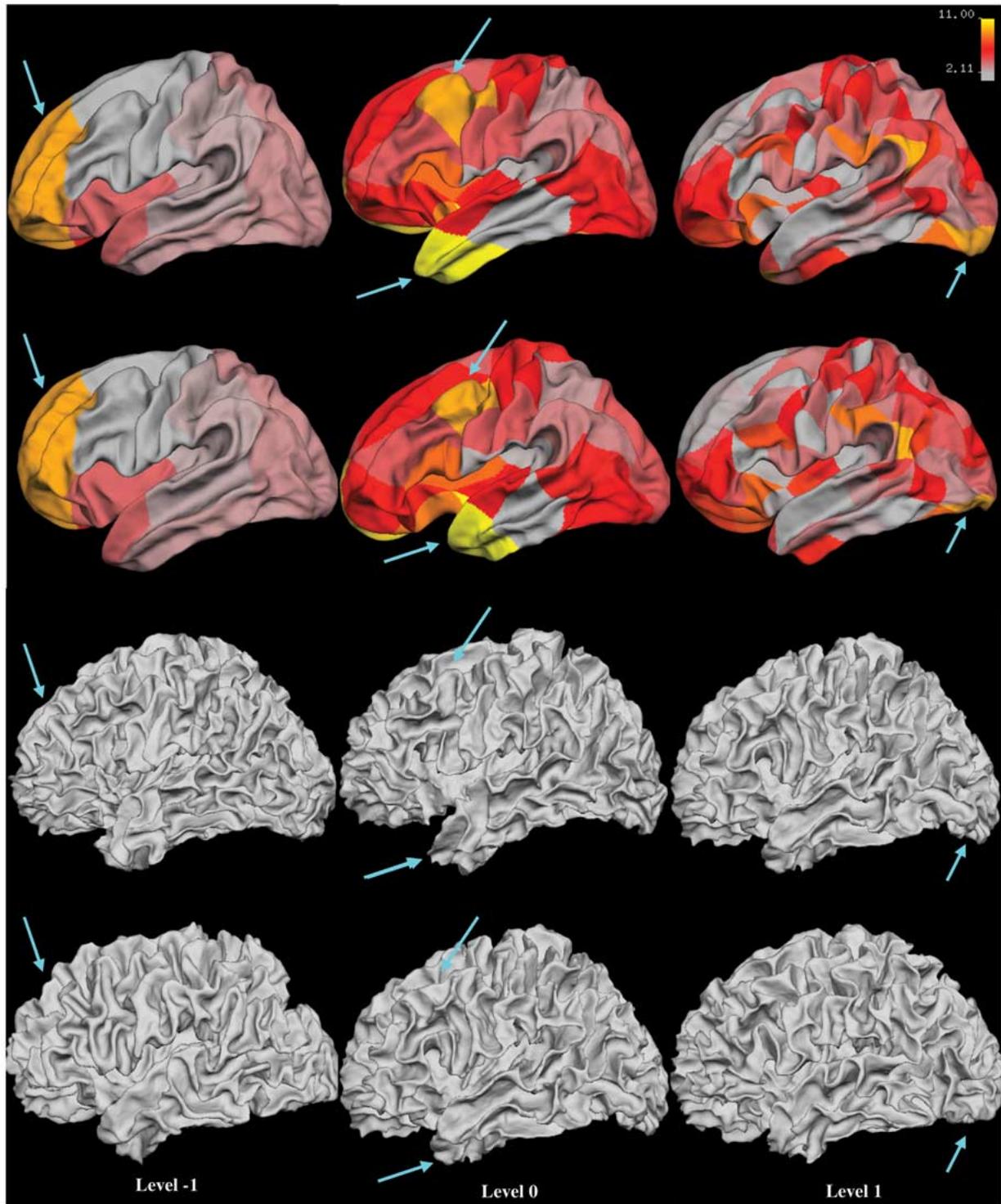
Fig. 5.   Principal surfaces at level −1 to 1 and real cortical surfaces. Top 2 rows: the synthetic surfaces representing the ∓3σ variations (ordered in top-down direction) of the first principal component at level −1, 0 and 1; Colormap shows the spatial coverage and $l_2$ norm of each wavelet coefficients in the first principal component; The color of each point on the cortical surface is assigned by the $l_2$ norm of the coefficient whose center of support is the closest; At the same level, the higher the intensity is, the larger the shape variation is across the entire dataset. Note that one cannot directly compare the colors at different levels, since the wavelet functions are not normalized across levels. (This normalization would lead to a wild difference between the wavelet coefficients across levels, masking the fine details in visualization.) Bottom 2 rows: Corresponding real surfaces that have the largest positive and negative projections on the first eigenvector (real ∓3σ surfaces). A set of arrows are used to point to the regions that vary the most across subjects at each level.

we are able to study the development of cortical folding at different spatial scales. Especially, the fitted Gompertz curves of wavelet power at lower levels demonstrate the development of primary folds since the lower level wavelet coefficients encode the shape variations of the low-resolution approximations of cortical surfaces (as shown in Fig. 3), and the curves at higher levels demonstrate the development of secondary and tertiary folds as they quantify shape variations at smaller spatial scales.
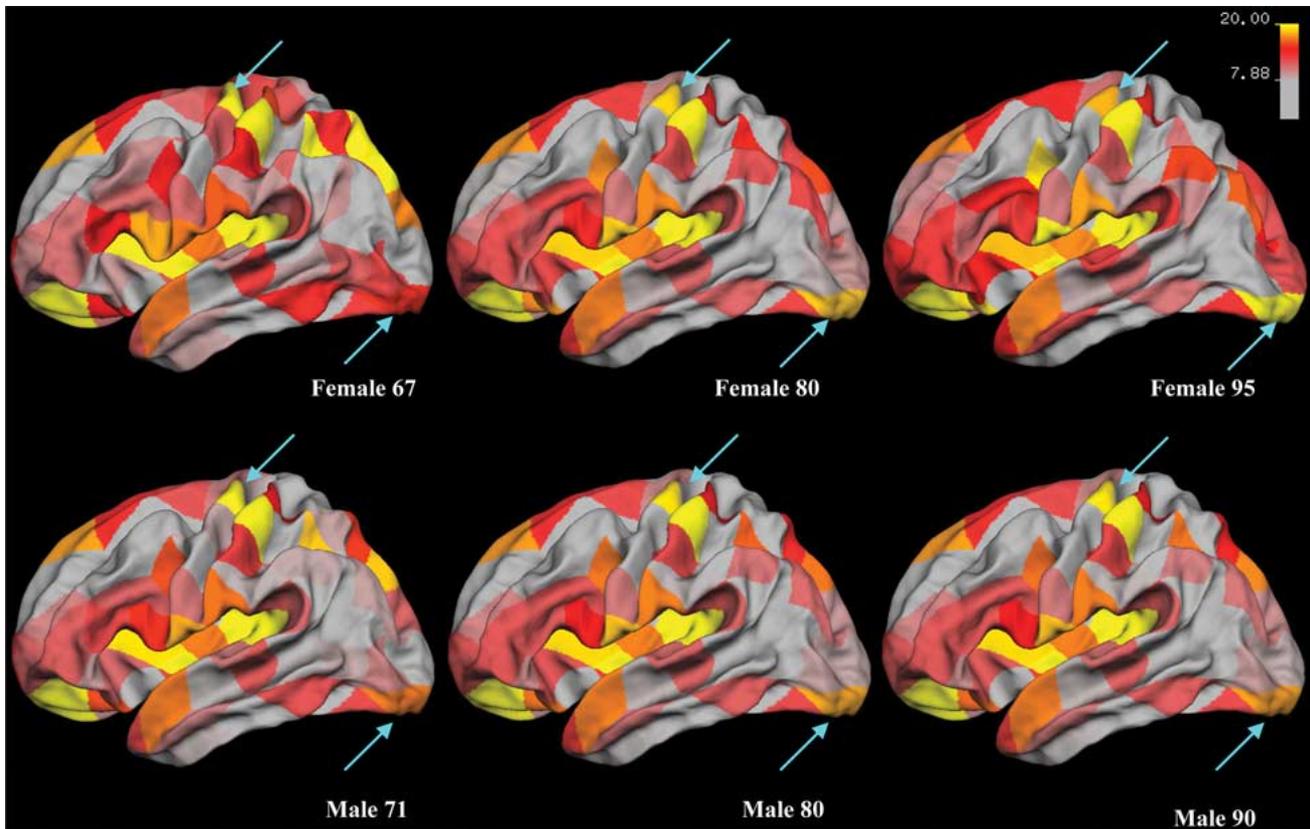
Fig. 6. Reconstructed surfaces for female (top row) and male (bottom row) subjects of different ages using projected wavelet coefficients on the set of principal components that represent 98% of the total variance at level 1. Colormap encodes the $l_2$ norm, location and support region of the projected coefficients of each subject. An arrow is pointed to a region of decreasing intensities with age on the precentral sulcus in the female group; another arrow is pointed to a region of increasing intensities with age on the occipital lobe; visual inspection identifies the narrowing of central sulcus and elongation of occipital lobe; neither these intensity changes nor shape changes are salient in the male group.

For the purpose of comparing wavelet powers across levels, they are normalized so that the $L^2$-norm of wavelet basis function is unity at all levels. As a result, the estimated parameters $g_1$ (the amount of folding), $g_2$ (speed of folding development), and $g_3$ (age of maximum folding development) can be used to quantitatively compare the folding developments across levels.

The results of this study are shown in Fig. 7 and Table II. First of all, Table II shows that the $R^2$ values are higher than 0.7 at most of the frequency levels in both hemispheres, indicating good model fitting results. As an example, the fitted Gompertz curves overlaid on the original data at level 0 in both hemispheres are shown in Fig. 7(a) and (b), with the vertical red line indicating the estimated age of maximum folding development (a logarithmic scale (base 10) is used for the horizontal axis). Secondly, Table II shows the estimated parameter $g_1$, whose value at different level encodes the amount of shape variations at the corresponding resolution. A similar value of $g_1$ within the 90% confidence intervals of the left and right hemisphere shows that the amounts of folds are equivalent on the left and right WM surfaces. Furthermore, Table II shows that the estimated speed of folding development increases from low– to high-frequency levels (levels 0–3) on both left- and right-hand sides, and the estimated age of estimated folding development increases monotonously with frequency level from approximately 29–33 weeks. The calculated 90% confidence intervals show that the estimated fastest development ages and speeds are significantly

different across frequency levels. These results indicate that the lower frequency folding such as the primary folds develop earlier and slower than the higher frequency folding such as the secondary and tertiary folds. At last, the comparison of left and right hemispheres shows that the speed of folding development is higher, but the age of fastest development is the same in the right hemisphere at levels 0–4, suggesting that folds develop simultaneously but faster on the right-hand side. To demonstrate these results, part of the fitted Gompertz curves including only the 8 newborns at different levels are plotted together in Fig. 7(c) and (d), where the red vertical lines indicate the maximum development ages estimated for folds of large and smaller scales. Although the data of three older children were not shown in Fig. 7(c) and (d), all 11 data points are used in the actual model fitting. A logarithmic scale (base 10) is used for the vertical axis for better visualization.

The cortical folding development model was then fitted to each one of the wavelet coefficients across 11 subjects. Unlike the study of wavelet power, this approach allows us to discover not only when, but also where the folding of the cortical surface occurs at different spatial scales. Although each of the $x$, $y$ and $z$ components of a wavelet coefficient was fitted to the model, only the estimated parameters $\{g_i\}_{i=1,...3}$ of the component with the largest $R^2$ value are used to demonstrate the development of folding in the support region of this wavelet function. The estimated folding development speed and age of the set of wavelet
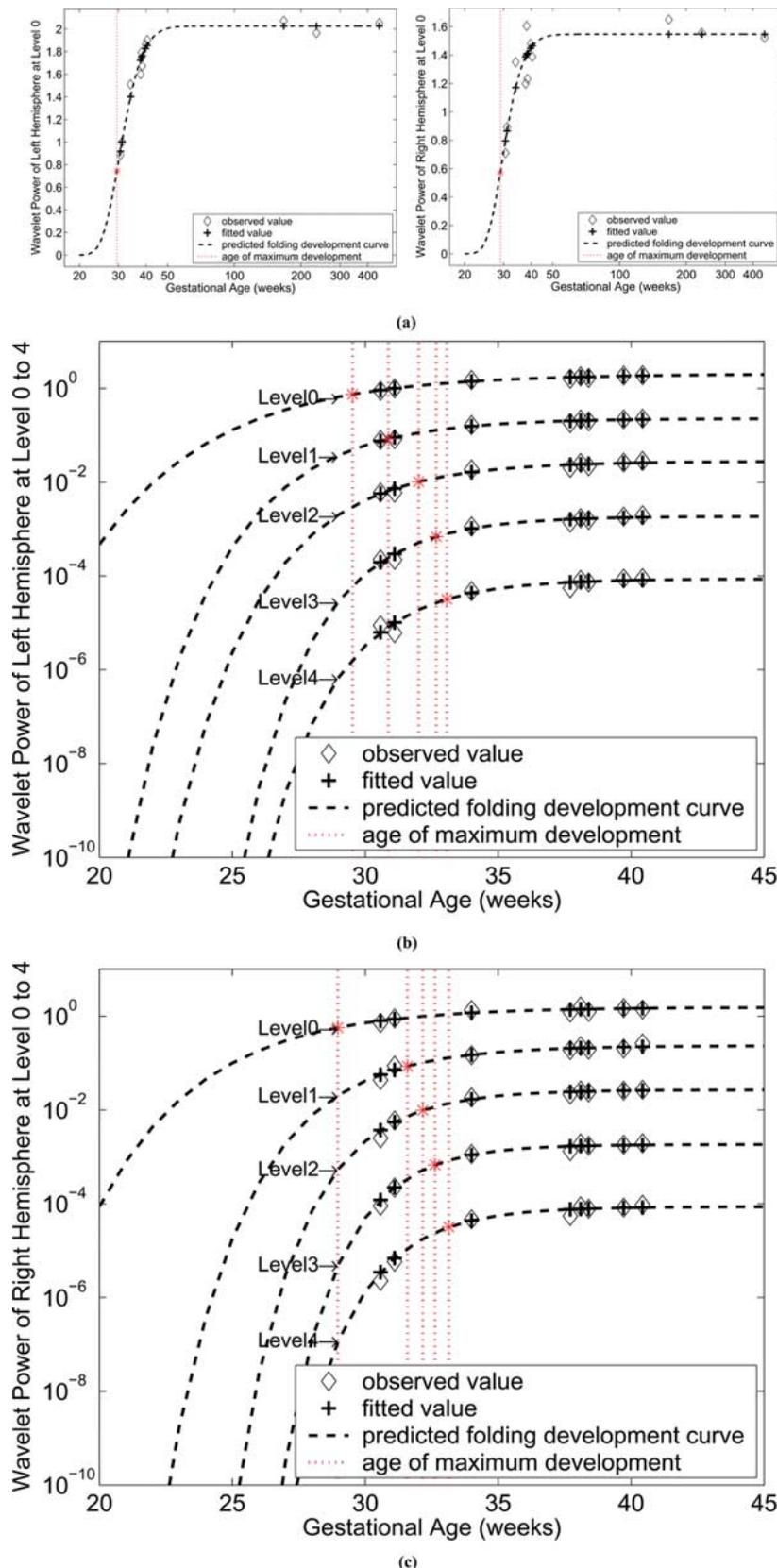
Fig. 7.   Predicted cortical folding development curves using wavelets power of the left and right hemispheres. (a) The predicted curves at level 0 only of left and right hemispheres (a logarithmic scale (base 10) is used for the horizontal from 0 to 450 weeks to include all the 11 data points used in the model fitting; vertical is the wavelet power normalized with regard to the wavelet basis function). (b) The predictive curves from frequency levels 0–4 of the left hemisphere (horizontal is the actual age from 20–45 weeks; vertical is the wavelet power with logarithmic (base 10) scale). (c) The predictive curves from frequency levels 0 to 4 of the right hemisphere.

coefficients with $R^2 > 0.5$ are plotted on the youngest newborn WM surface with a colormap showing the estimated parame-

ters, and the location and support regions of their corresponding wavelet basis functions. The figures of maximum development

| | Level | $g_1$ | Development Speed (1/week) | Maximum Development Age (week) | $R^2$ |
|---|---|---|---|---|---|
| | 0 | $2.03 \pm 0.20$ | $0.22 \pm 9.3\text{e-}4$ | $29.52 \pm 0.221$ | 0.82 |
| | 1 | $2.27\text{e-}1 \pm 1.59\text{e-}2$ | $0.31 \pm 3.1\text{e-}3$ | $30.87 \pm 0.128$ | 0.81 |
| Left | 2 | $2.78\text{e-}2 \pm 1.90\text{e-}3$ | $0.32 \pm 4.2\text{e-}3$ | $32.01 \pm 0.229$ | 0.78 |
| | 3 | $1.87\text{e-}3 \pm 3.20\text{e-}5$ | $0.39 \pm 6.2\text{e-}3$ | $32.67 \pm 0.201$ | 0.81 |
| | 4 | $8.70\text{e-}5 \pm 1.00\text{e-}6$ | $0.39 \pm 1.1\text{e-}2$ | $33.06 \pm 0.417$ | 0.75 |
| | 0 | $1.55 \pm 0.72$ | $0.25 \pm 1.0\text{e-}2$ | $28.97 \pm 1.453$ | 0.57 |
| | 1 | $2.33\text{e-}1 \pm 5.91\text{e-}2$ | $0.34 \pm 1.1\text{e-}2$ | $31.59 \pm 0.364$ | 0.70 |
| Right | 2 | $2.67\text{e-}2 \pm 1.10\text{e-}3$ | $0.43 \pm 7.0\text{e-}3$ | $32.16 \pm 0.119$ | 0.83 |
| | 3 | $1.83\text{e-}3 \pm 3.70\text{e-}5$ | $0.49 \pm 1.8\text{e-}2$ | $32.62 \pm 0.239$ | 0.78 |
| | 4 | $8.60\text{e-}5 \pm 1.00\text{e-}6$ | $0.45 \pm 2.7\text{e-}2$ | $33.14 \pm 0.467$ | 0.73 |

ages also show an earlier development of large-scale folds in both left and right hemispheres. The comparison of the age and speed figures of the same region at the same level shows that in general the earlier a region develops (darker blue in the maximum development age colormap), the slower the development speed is (more red and less yellow in the speed colormap).

Furthermore, we can determine where and how fast the folding occurs on the WM surface at different ages. For this purpose, we first cluster wavelet coefficients into three age intervals (0, 33), (33, 38), and (38, ∞) weeks by using their estimated maximum development ages. Then the estimated development speed of the wavelet coefficients in each age interval is plotted as colormap in the support regions of the corresponding wavelet basis functions on a representative brain (30, 34, and 40 weeks old newborns for the three age intervals respectively), as shown in Fig. 9. The colormap on each surface encodes the estimated speed of the regions that develop the fastest within the specific age interval. Visual inspection shows that most regions that develop at younger ages are of larger scales, and regions that develop at older ages are of smaller scales. This observation further demonstrates the earlier development of larger primary folds followed by the development of secondary and tertiary folds at successively higher scales.

Modeling the cortical folding in the wavelet domain allows us to evaluate the primary folds and smaller scale folds separately in the brain development. Studies of both wavelet power and individual coefficients show larger scales of folding development at younger ages with slower speeds. Moreover, the individual wavelet study quantifies the localized folding developments in different cortical regions and age ranges, and further demonstrates the positive correlation between estimated maximum development age and speed at the same level.

However, the correlation between wavelet coefficients at different levels should be taken into account when assessing the results. In this paper, most of the detected brain regions of fast folding development do not overlap with each other across levels, as shown in Figs. 8 and 9. In the overlapped regions, the correlations between overlapping wavelet bases are fairly small. For example, a smaller scale fold on the temporal lobe, corresponding to a wavelet coefficient at level 2, is detected to be in the region of a larger scale fold, corresponding to a wavelet coefficient at level 0, as seen in the upper left image in Fig. 9. The correlation coefficient of the two corresponding wavelet basis functions is calcu-

lated as 0.031, showing a weak correlation of the folding development detected in these two regions. Therefore, the biorthogonality property of the wavelet bases has minor effects on our study of cortical folding development by using wavelet coefficients at different levels separately.

Another limitation is the observation that many regions on the WM surface do not fit well to the Gompertz model. Other models are currently being explored to account for these regions.

## IV. CONCLUSION

The spherical wavelet transformation was demonstrated to be able to accurately and efficiently detect the locations and spatial scales of shape variations. The use of wavelet coefficients in detecting and visualizing patterns of cortical surface variation shows promising results in a nondemented aged population. The study of cortical surface folding development in newborns also demonstrated the power of wavelets in analyzing the underlying function locally in both the space and the frequency domain. The regularized Gompertz function applied to this population was shown to provide a powerful model for the observed folding development as characterized by the wavelet coefficients, and allowed the generation of maps revealing the temporal ordering of the development of large scale and progressively finer scale folds. Future work includes employing more sophisticated statistical tools and extending the wavelet analysis to other neuroanatomical structures.

## APPENDIX

### A. Fast Spherical Wavelets Transformation

In this paper, we used the interpolation scaling function defined as

$$\varphi_{j,k}(v_{j,k'}) = \delta_{k-k'} \quad for \quad k, k' \in K(j) \tag{A-1}$$

where $v_{j,k}$ denotes the $k$th vertex at subdivision level $j (k \in K(j))$. The same notation is used here as in the main text so that $K(j)$ and $M(j)$ denote all vertices on the mesh before and after the $j$th subdivision. As a result, the scaling coefficients at level $j$ of a function $f$ interpolated on the $j^{\text{th}}$ order icosahedron are the values of this function at each vertex, i.e.,

$$\lambda_{j,k} = f(v_{j,k}), \ k \in K(j). \tag{A-2}$$

A butterfly subdivision scheme is used such that the value of a scaling coefficient at level $j + 1$ can be found as

$$\lambda_{j+1,m} = 1/2 \sum_{k \in A(j,m)} \lambda_{j,k} + 1/8 \sum_{k \in B(j,m)} \lambda_{j,k} \\ - 1/16 \sum_{k \in C(j,m)} \lambda_{j,k} \tag{A-3}$$

where $A(j,m)$, $B(j,m)$, and $C(j,m)$ are local neighbors of vertex $v_{j+1,m}$ as shown in Fig. (A-1). The butterfly procedure is used to generate smooth wavelet functions.

Next, a lifting algorithm is used to construct wavelet function as

$$\psi_{j,m} = \varphi_{j+1,m} - \sum_{k \in A(j,m)} s_{j,k,m} \varphi_{j,k} \tag{A-4}$$
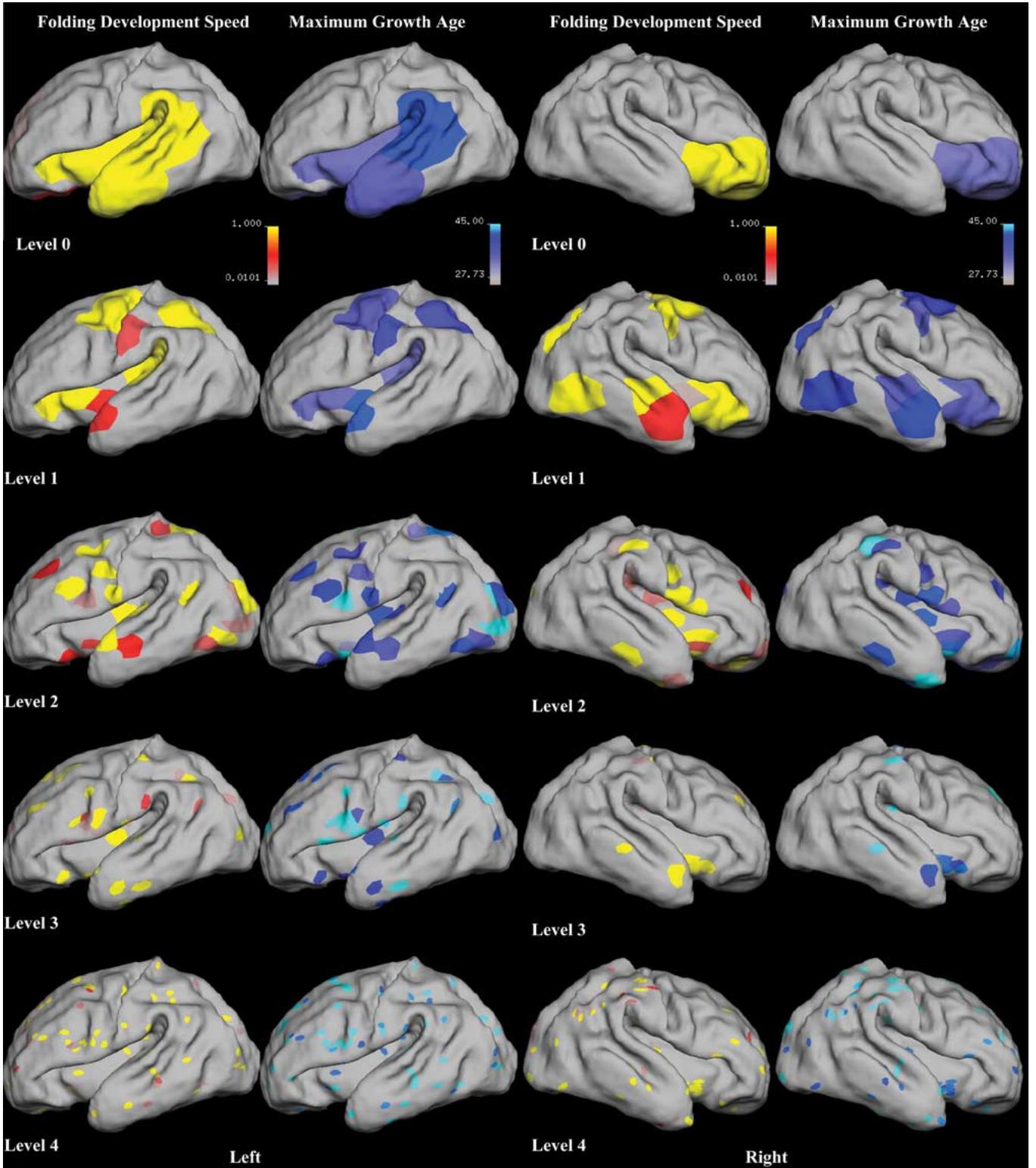
Fig. 8. The predicted folding development speed and maximum development ages for the left and right hemispheres using individual wavelets at level 0–4 with colormaps encoding the magnitude of the estimated development speed (1/week) and age of maximum development (weeks) of wavelet coefficients in the support regions of their corresponding wavelet basis functions. For points in the overlapped regions of two or more wavelet basis functions, the estimated age and speed of the closest wavelet function is assigned. Column 1: predicted folding development speed in the left hemisphere from level 0–4 (top-down); column 2: predicted age of maximum folding development in the left hemisphere from level 0–4 (top-down); column 3: folding development speed in the right hemisphere from level 0–4; column 4: age of maximum folding development in the right hemisphere from level 0–4.

where $s_{j,k,m} = I_{j+1,m}/I_{j,k}$ with $I_{j,k} = \int \varphi_{j,k} d\omega$. This ensures that the constructed wavelet function has a vanishing integral, i.e. one vanishing moment. The wavelet function constructed in this way has a local support in frequency because its value is vanishing in both the high-frequency due to the smoothness and low frequency due to the vanishing integral.
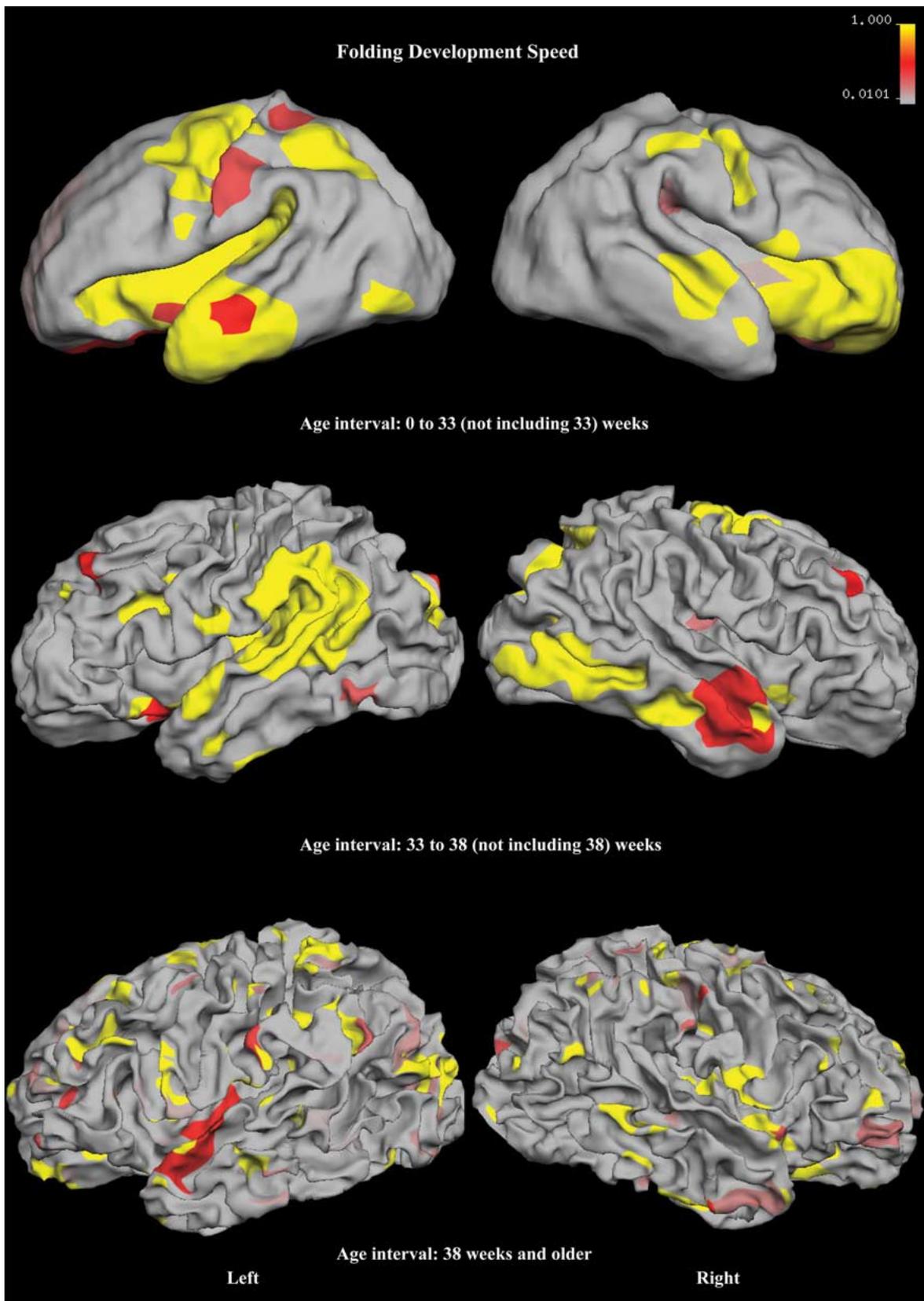
Fig. 9. Estimated folding development speed of the regions that develop the fastest during age 0–33 weeks (first row), 33–38 weeks (second row), and 38 weeks and older (third row). Colormap encodes the magnitude of the estimated development speed (1/week) in the support regions of their corresponding wavelet basis functions. For points in the overlapped regions of two or more wavelet basis functions, the estimated speed of the highest level or closest wavelet function is assigned.

In this paper, the coordinate function is interpolated onto the seventh-order icosahedron. The resulting scaling coefficients at the highest level are therefore the values of the interpolated coordinate function at each vertex on the icosahedron
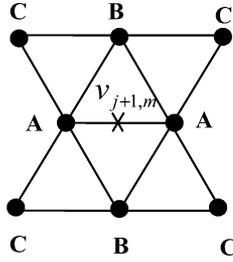
Fig. 10. Local neighborhoods for the Butterfly scheme, where vertex sets A, B, and C are used in determining the new value at the center edge midpoint.

at subdivision level $= 7$. The wavelet coefficients can be calculated in two steps (Analysis step I and II) recursively as shown below.

Analysis Step I: Calculate the $\gamma_{j,m}$ as follows:

$$\forall m \in M(j) : \gamma_{j,m} := \lambda_{j+1,m} - 1/2 \sum_{k \in A(j,m)} \lambda_{j,k}$$
$$-1/8 \sum_{k \in B(j,m)} \lambda_{j,k} + 1/16 \sum_{k \in C(j,m)} \lambda_{j,k}.$$

Analysis Step II: Calculate the $\lambda_{j,k}$ using the $\gamma_{j,m}$ from step I:

$$\forall k \in K(j) : \lambda_{j,k} = \lambda_{j+1,k}$$
$$\forall m \in M(j) : \forall k \in A(j,m) : \lambda_{j,k} + = s_{j,k,m}\gamma_{j,m}.$$

The inverse transformation can be implemented in two steps (Synthesis step I and II) as well.

Synthesis Step I: Calculate the $\lambda_{j+1,k}$:

$$\forall k \in K(j) : \lambda_{j+1,k} = \lambda_{j,k}$$
$$\forall m \in M(j) : \forall k \in A(j,m) : \lambda_{j+1,k} - = s_{j,k,m}\gamma_{j,m}.$$

Synthesis Step II: Calculate $\lambda_{j+1,m}$ using the $\lambda_{j+1,k}$ from step I:

$$\forall m \in M(j) : \lambda_{j+1,m} := \gamma_{j,m} + 1/2 \sum_{k \in A(j,m)} \lambda_{j+1,k}$$
$$+1/8 \sum_{k \in B(j,m)} \lambda_{j+1,k} - 1/16 \sum_{k \in C(j,m)} \lambda_{j+1,k}.$$

## REFERENCES

[1] F. L. Bookstein, *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[2] I. Dryden and K. Mardia, "Multivariate shape analysis," *Sankhya*, vol. 55, pp. 460–480, 1993.

[3] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models—Their training and application," *Comput. Vis. Image Understanding*, vol. 61, pp. 38–59, 1995.

[4] C. Brechbüler, G. Gerig, and O. Kubler, "Parametrization of closed surfaces for 3-D shape description," *Comput. Vis. Image Understanding*, vol. 61, pp. 154–170, 1995.

[5] A. Kelemen, G. Székely, and G. Gerig, "Elastic model-based segmentation of 3D neuroradiological data sets," *IEEE Trans. Med. Imag.*, vol. 18, pp. 828–839, 1999.

[6] M. Styner, G. Gerig, J. Lieberman, D. Jones, and D. Weinberger, "Statistical shape analysis of neuroanatomical structures based on medial models," *Med. Image Anal.*, vol. 7, pp. 207–220, 2003.

[7] L. Shen, J. Ford, F. Makedon, and A. Saykin, "Hippocampal shape analysis surface-based representation and classification," *Proc. SPIE Medical Imaging*, 2003.

[8] M. K. Chung, S. Wang, K. M. Dalton, R. J. Davidson, S. Robbins, and A. C. Evans, "Tensor-based cortical morphometry via weighted spherical harmonic representation," presented at the MMBIA Conf, 2006, unpublished.

[9] G. Christensen, R. Rabbitt, and M. Miller, "3D brain mapping using a deformable neuroanatomy," *Physics in Medicine and Biology*, vol. 39, pp. 609–618, 1994.

[10] C. Davatzikos, M. Vaillant, S. Resnick, J. Prince, S. Letovsky, and R. Bryan, "A computerized method for morphological analysis of the corpus callosum," *J. Comput. Assist. Tomogr.*, vol. 20, pp. 88–97, 1996.

[11] S. C. Joshi, M. I. Miller, and U. Grenander, "On the geometry and shape of brain sub-manifolds," *Int. J. Pattern Recognit. Artif. Intell. (Special Issue Process. MR Images Hum. Brain)*, vol. 11, no. 8, pp. 1317–1343, 1997.

[12] P. Thompson, J. Giedd, R. Woods, D. MacDonald, A. Evans, and A. Toga, "Growth patterns in the developing brain detected by using continuum machanical tensor maps," *Nature*, pp. 190–193, 2000.

[13] T. Blum, "A transformation for extracting new descriptors of shape," in *Models for the Perception of Speech and Visual Form*. Cambridge: MIT Press, 1967, pp. 362–380.

[14] S. Pizer, D. Fritsch, P. Yushkevich, V. Johnson, and E. Chaney, "Segmentation, registration, and measurement of shape variation via image object shape," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 851–865, Oct. 1999.

[15] P. Golland, W. Grimson, and R. Kikinis, "Statistical shape analysis using fixed topology skeletons: Corpus callosum study," in *Proc. Inf. Process. Med. Imag.*, 1999, pp. 382–388.

[16] C. Davatzikos, X. Tao, and D. Shen, "Hierarchical active shape models, using the wavelet transform," *IEEE Trans. Med. Imag.*, vol. 22, no. 3, pp. 414–423, Mar. 2003.

[17] D. Nain, S. Haker, A. Bobick, and A. Tannenbaum, "Multiscale 3D shape analysis using spherical wavelets," in *Proc. MICCAI*, Oct. 2005, pp. 459–467.

[18] D. Nain, S. Haker, A. Bobbick, and A. Tannenbaum, "Shape-driven surface segmentation using spherical wavelets," in *Proc. MICCAI*, 2006, pp. 17–24.

[19] C. G. Broyden, "An alternative derivation of simplex method," *J. Inst. Math. Appl.*, vol. 6, pp. 76–90, 1970.

[20] A. M. Dale, B. Fischl, and M. I. Sereno, "Cortical surface-based analysis I: Segmentation and surface reconstruction," *Neuroimage*, vol. 9, pp. 179–194, 1999.

[21] B. Fischl, M. I. Sereno, and A. M. Dale, "Cortical surface-based analysis II: Inflation, flattening, and a surface-based coordinate system," *Neuroimage*, vol. 9, pp. 195–207, 1999.

[22] B. Fischl, A. M. Dale, M. I. Sereno, R. B. H. Tootell, and B. R. Rosen, "A coordinate system for the cortical surface," *NeuroImage*, vol. 7, p. S740, 1998.

[23] L. Shen and M. K. Chung, "Large-scale modeling of parametric surfaces using spherical harmonics," presented at the 3rd Int. Symp. 3D Data Processing, Visualization and Transmission (IEEE 3DPVT), Chapel Hill, NC, Jun. 14–16, 2006, unpublished.

[24] I. Daubechies, *Ten lectures on wavelets*, ser. Appl. Math. Philadelphia, PA: SIAM, 1992, vol. 61.

[25] P. Schröder and W. Sweldens, "Spherical wavelets: Efficiently representing functions on a sphere," in *Proc. SIGGRAPH '95*, 1995, pp. 161–172.

[26] A. Cohen, I. Daubechies, and J. Faveau, "Biorthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 45, pp. 485–560, 1992.

[27] E. J. Stollnitz, T. D. DeRose, and D. H. Salesin, *Wavelets for Computer Graphics: Theory and Applications*. San Francisco, CA: Morgan Kaufmann, 1996.

[28] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Training models of shape from sets of examples," in *Proc. Br. Machine Vision Conf.*, 1992, pp. 9–18, Springer-Verlag.

[29] G. Székely *et al.*, "Segmentation of 2D and 3D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface models," *Med. Image Anal.*, vol. 1, pp. 19–34, 1996.

[30] M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *Proc. CVPR'2000*, 2000, pp. 316–323.

[31] D. Shen and C. Davatzikos, "An adaptive-focus deformable model using statistical and geometric information," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, pp. 906–913, 2000.

[32] V. S. J. Caviness, "Mechanical model of brain convolutional development," *Science*, vol. 189, pp. 18–21, 1975.

[33] D. C. V. Essen, "A tension-based theory of morphogenesis and compact wiring in the central nervous system," *Nature*, vol. 385, pp. 313–318, 1997.

[34] E. P. Virene, "Reliability growth and its upper limit," in *Proc. 1968 Annu. Symp. Reliability*, 1968, pp. 265–270.

[35] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Oxford Umiv. Press, 1995.

[36] T. P. Minka, Using Lower Bounds to Approximate Integrals Carnegie Mellon Univ., Pittsburgh, PA, Tech Rep., Jul. 2001.